Journal Code:				Article ID					Dispatch 10.12.13	CE: Patrimonio, Cheryl F.
S	J	O	S	1	2	6	9		No. of Pages: 27	ME:

Scandinavian Journal of Statistics
doi: 10.1111/sjos.1269

Asymptotic Behavior of Conditional Least Squares Estimators for Unstable Integer-valued Autoregressive of Order 2 Models

MÁTYÁS BARCZY and MÁRTON ISPÁNY
Faculty of Informatics, University of Debrecen
GYULA PAP
Bolyai Institute, University of Szeged

ABSTRACT. In this paper, the asymptotic behavior of the conditional least squares estimators of the autoregressive parameters, of the mean of the innovations, and of the stability parameter for unstable integer-valued autoregressive processes of order 2 is described. The limit distributions and the scaling factors are different according to the following three cases: (i) decomposable, (ii) indecomposable but not positively regular, and (iii) positively regular models.

Key words: branching process with immigration, Bessel process, conditional least squares estimator, martingale, unstable INAR(p) process

1. Introduction

The theory and practice of statistical inference for integer-valued time series models are rapidly developing and important topics of the modern theory of statistics; see, for example, Steutel & van Harn (1979) and Weiß (2008).

Among the most successful integer-valued time series models proposed in the literature, we mention the integer-valued autoregressive model of order p (INAR(p)). Such a model was first introduced by Alzaid & Al-Osh (1990). Another definition of INAR(p) processes was proposed independently by Du & Li (1991) and by Gauthier & Latour (1994) and Latour (1998). In Du and Li's approach, the autocorrelation structure of an INAR(p) process is the same as that of an autoregressive model of order p (AR(p)) process, and we follow this setup. In Barczy *et al.* (2011), we investigated the asymptotic behavior of unstable INAR(p) processes, that is, when the characteristic polynomial has a unit root. Under some natural assumptions, we proved that the sequence of appropriately scaled random step functions formed from an unstable INAR(p) process converges weakly towards a squared Bessel process. This limit process is a continuous time branching process with immigration also known as the Cox–Ingersoll–Ross process. AQ1

Parameter estimation for INAR(p) models has a long history. Franke & Seligmann (1993) analyzed the conditional maximum likelihood estimator of some parameters (including the autoregressive parameters) for stable switching INAR(1) models with Poisson innovations. Du & Li (1991), in Theorem 4.2, proved asymptotic normality of the conditional least squares (CLS) estimator of the autoregressive parameters for stable INAR(p) models; see also Proposition 6.1 of Latour (1998); Brännäs & Hellström (2001) considered generalized method of moment estimation. Ispány *et al.* (2003a), (2003b) derived asymptotic inference for nearly unstable INAR(1) models, which has been refined by Drost *et al.* (2009) later. In Ispány *et al.* (2003a), the mean of the innovations was supposed to be known, whereas in Ispány *et al.*

(2003b), both the autoregressive parameter and the mean of the innovations have been estimated jointly. Drost *et al.* (2008) studied asymptotically efficient estimation of the parameters for stable INAR(p) models. The stability parameter $\varrho := \alpha_1 + \cdots + \alpha_p$ of an INAR(p) model with autoregressive parameters $(\alpha_1, \ldots, \alpha_p)$ has not been treated yet, but the asymptotic behavior of the CLS estimator of this stability parameter is well investigated in case of unstable AR(p) processes; see the unit root tests, for example, in Section 17, Table 17.3, Case 1 of Hamilton (1994). To the best of our knowledge, unit root tests for general INAR(p) models are not known, and from this point of view, studying unstable INAR(p) models is an important preliminary task. In this paper, the asymptotic behavior of the CLS estimators of the autoregressive and stability parameters together with the mean of the innovations for unstable INAR(2) models is described (see our main results in Section 2), which can be considered as a first step of examining this question for general unstable INAR(p) processes and more generally for critical multitype branching processes. We call the attention that in case of unstable INAR(2) processes, new types of limit distributions occur (Theorem 2.1) compared with those of unstable AR(p) processes.

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers, real numbers, and non-negative real numbers, respectively. Every random variable will be defined on a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition 1.1. *Let $(\varepsilon_k)_{k\in\mathbb{N}}$ be an independent and identically distributed (i.i.d.) sequence of non-negative integer-valued random variables, and let $(\alpha, \beta) \in [0,1]^2$. An INAR(2) model with autoregressive parameters (α, β) and innovations $(\varepsilon_k)_{k\in\mathbb{N}}$ is a stochastic process $(X_k)_{k\geq -1}$ given by*

$$X_k = \sum_{j=1}^{X_{k-1}} \xi_{k,j} + \sum_{j=1}^{X_{k-2}} \eta_{k,j} + \varepsilon_k, \qquad k \in \mathbb{N}, \tag{1.1}$$

where for all $k \in \mathbb{N}$, $(\xi_{k,j})_{j\in\mathbb{N}}$ and $(\eta_{k,j})_{j\in\mathbb{N}}$ are sequences of i.i.d. Bernoulli random variables with mean α and β, respectively, such that these sequences are mutually independent and independent of the sequence $(\varepsilon_k)_{k\in\mathbb{N}}$, and X_0 and X_{-1} are non-negative integer-valued random variables independent of the sequences $(\xi_{k,j})_{j\in\mathbb{N}}$, $(\eta_{k,j})_{j\in\mathbb{N}}$, $k \in \mathbb{N}$, and $(\varepsilon_k)_{k\in\mathbb{N}}$.

An INAR(2) model (1.1) can also be written in the form $X_k = \alpha \circ X_{k-1} + \beta \circ X_{k-2} + \varepsilon_k$, $k \in \mathbb{N}$, using the binomial thinning operator $\circ$ due to Steutel & van Harn (1979).

For the sake of simplicity, we consider a zero-start INAR(2) process; that is, we suppose $X_0 = X_{-1} = 0$. The general case of non-zero initial values may be handled in a similar way.

In the sequel, we always assume $\mathbb{E}(\varepsilon_1^2) < \infty$. Let us denote the mean and variance of ε_1 by μ and σ^2, respectively. Further, we assume $\mu > 0$, otherwise $X_k = 0$ for all $k \in \mathbb{N}$.

On the basis of the asymptotic behavior of $\mathbb{E}(X_k)$ as $k \to \infty$ described in Proposition 2.6 of Barczy *et al.* (2011), we distinguish three types of INAR(2) models. This asymptotic behavior is determined by the spectral radius r of the matrix

$$A := \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix}.$$

The case $r < 1$, when $\mathbb{E}(X_k)$ converges to a finite limit as $k \to \infty$, is called *stable* or *asymptotically stationary*, whereas the cases $r = 1$, when $\mathbb{E}(X_k)$ tends linearly to ∞, and $r > 1$, when $\mathbb{E}(X_k)$ converges to ∞ with an exponential rate, are called *unstable* and *explosive*, respectively. It is easy to check that $r < 1$, $r = 1$, and $r > 1$ are equivalent with $\varrho < 1$, $\varrho = 1$, and

$\varrho > 1$, respectively, where $\varrho := \alpha + \beta$ is called the *stability parameter*; see Proposition 2.2 of Barczy *et al.* (2011).

We also note that an INAR(2) process can be considered as a special 2-type branching process with immigration. Namely, by (1.1),

$$\begin{bmatrix} X_k \\ X_{k-1} \end{bmatrix} = \sum_{j=1}^{X_{k-1}} \begin{bmatrix} \xi_{k,j} \\ 1 \end{bmatrix} + \sum_{j=1}^{X_{k-2}} \begin{bmatrix} \eta_{k,j} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_k \\ 0 \end{bmatrix}, \qquad k \in \mathbb{N},$$

and hence the so-called mean matrix of an INAR(2) process with autoregressive parameters (α, β) (considered as a 2-type branching process) is nothing else but A. This process is called *positively regular*, if there is a positive integer $k \in \mathbb{N}$ such that the entries of A^k are positive; see Kesten & Stigum (1966a), which is equivalent with $\alpha > 0$ and $\beta > 0$. The model is called *decomposable*, if the matrix A is decomposable; see Kesten & Stigum (1967), which is equivalent with $\beta = 0$. If $\alpha = 0$ and $\beta > 0$, then the process is *indecomposable but not positively regular*; see Kesten & Stigum (1966b). For more details of this classification, see Appendix A of Barczy *et al.* (2012).

Next, we give an overview of the structure of the paper. Section 2 contains our main results; see Theorem 2.1 for unstable and positively regular INAR(2) processes, Theorem 2.5 for unstable and decomposable INAR(2) processes, and Theorem 2.7 for unstable, indecomposable but not positively regular ones. In order to highlight our main results, the preliminaries and (technical) details on CLS estimators are presented only after our main results; see Section 3. In Theorems 4.1, 4.2, and 4.3 of Section 4, we present joint asymptotic behaviors of the building blocks of the CLS estimators (according to the aforementioned three cases), and by applying a version of the continuous mapping theorem (which is formulated for completeness in Appendix Appendix B), we show how one can derive Theorems 2.1, 2.5, and 2.7 using these theorems. Section 5 is devoted to the proof of Theorem 4.1, which is based on Lemma 5.1 and Theorem 5.1. Because of its length, the proof of Theorem 5.1 is given separately in Section 6. Sections 7 and 8 are devoted to the proofs of Theorems 4.2 and 4.3, respectively. In Appendix Appendix A, we present estimates for the moments of the processes involved; these estimates are used throughout the paper. In Appendix Appendix C, we recall a result about convergence of random step processes noting that the proof of Theorem 5.1 is based on this result. For detailed proofs, see Barczy *et al.* (2012).

2. Main results

In what follows, we always assume $\varrho = \alpha + \beta = 1$; that is, the process $(X_k)_{k \geq -1}$ is unstable.

For each $n \in \mathbb{N}$, any CLS estimator $\left(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n)\right)$ of the autoregressive parameters (α, β) and of the mean μ of the innovations based on a sample $\boldsymbol{X}_n := (X_1, \ldots, X_n)$ has the form

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \left(\sum_{k=1}^{n} \begin{bmatrix} X_{k-1}^2 & X_{k-1}X_{k-2} & X_{k-1} \\ X_{k-1}X_{k-2} & X_{k-2}^2 & X_{k-2} \\ X_{k-1} & X_{k-2} & 1 \end{bmatrix} \right)^{-1} \sum_{k=1}^{n} \begin{bmatrix} X_k X_{k-1} \\ X_k X_{k-2} \\ X_k \end{bmatrix},$$

on the set $\left\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\right\}$ with $\mathbb{P}\left(\sum_{k=1}^{n} X_{k-2}^2 > 0\right) \to 1$ as $n \to \infty$, see Proposition 3.1. Further, for each $n \in \mathbb{N}$, any CLS estimator of the stability parameter ϱ takes the form $\widehat{\varrho}_n(\boldsymbol{X}_n) = \widehat{\alpha}_n(\boldsymbol{X}_n) + \widehat{\beta}_n(\boldsymbol{X}_n)$ on the set $\left\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\right\}$; see Section 3.

Theorem 2.1. Let $(X_k)_{k\geq -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in (0,1)^2$ such that $\alpha + \beta = 1$ (hence, it is unstable and positively regular). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}\left(\varepsilon_1^8\right) < \infty$, and $\mu > 0$. Then

$$n\left(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1\right) \xrightarrow{\mathcal{L}} \frac{\sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t^{3/2}\,\mathrm{d}\mathcal{W}_t - [(1+\beta)\mathcal{X}_1 - \mu]\int_0^1 \mathcal{X}_t\,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2}, \tag{2.1}$$

$$\begin{bmatrix} n^{1/2}\left(\widehat{\alpha}_n(\boldsymbol{X}_n) - \alpha\right) \\ n^{1/2}\left(\widehat{\beta}_n(\boldsymbol{X}_n) - \beta\right) \end{bmatrix} \xrightarrow{\mathcal{L}} \sqrt{\alpha(1+\beta)}\,\frac{\int_0^1 \mathcal{X}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{X}_t\,\mathrm{d}t}\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \tag{2.2}$$

$$\widehat{\mu}_n(\boldsymbol{X}_n) - \mu \xrightarrow{\mathcal{L}} \frac{-\sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t\,\mathrm{d}t\int_0^1 \mathcal{X}_t^{3/2}\,\mathrm{d}\mathcal{W}_t + [(1+\beta)\mathcal{X}_1 - \mu]\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2}, \tag{2.3}$$

as $n \to \infty$, where $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution and $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is the unique strong solution of the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{X}_t = \frac{1}{1+\beta}\left(\mu\,\mathrm{d}t + \sqrt{2\alpha\beta\mathcal{X}_t^+}\,\mathrm{d}\mathcal{W}_t\right), \qquad t \in \mathbb{R}_+, \tag{2.4}$$

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$, $\left(\widetilde{\mathcal{W}}_t\right)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes and x^+ denotes the positive part of $x \in \mathbb{R}$.

Remark 2.2. The moment condition $\mathbb{E}\left(\varepsilon_1^8\right) < \infty$ in Theorem 2.1 seems to be too strong, but we call the attention that the process $(X_k)_{k\geq -1}$ can be considered as a heteroscedastic time series. Indeed, $X_k = \alpha X_{k-1} + \beta X_{k-2} + M_k + \mu$; see (3.3), and by (A.1), $\mathbb{E}\left(M_k^2 \mid \mathcal{F}_{k-1}\right) = \alpha(1-\alpha)X_{k-1} + \beta(1-\beta)X_{k-2} + \sigma^2$, $k \in \mathbb{N}$. That is why we think that the behavior of the process $(X_k)_{k\geq -1}$ is similar to generalized autoregressive conditional heteroscedasticity models, where, even in the stable case, high moment conditions are needed for convergence of AQ2 estimators such as the quasi-maximum likelihood estimator in Hall & Yao (2003) or the Whittle estimator in Mikosch & Straumann (2002).

Remark 2.3. The SDE (2.4) has a unique strong solution $\left(\mathcal{X}_t^{(x)}\right)_{t\in\mathbb{R}_+}$ for all initial values $\mathcal{X}_0^{(x)} = x \in \mathbb{R}$, and if the initial value $\mathcal{X}_0^{(x)} = x$ is non-negative, then $\mathcal{X}_t^{(x)}$ is non-negative for all $t \in \mathbb{R}_+$ with probability one; hence, $\mathcal{X}_t^+$ may be replaced by $\mathcal{X}_t$ under the square root in (2.4); see, for example, in Remark 3.3 of Barczy *et al.* (2011).

Remark 2.4. By Itô's formula and Remark 2.3, $\mathcal{M}_t := (1+\beta)\mathcal{X}_t - \mu t$, $t \in \mathbb{R}_+$, is the unique strong solution of the SDE

$$\mathrm{d}\mathcal{M}_t = \sqrt{\frac{2\alpha\beta}{1+\beta}\left(\mathcal{M}_t + \mu t\right)^+}\,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+, \tag{2.5}$$

with initial value $\mathcal{M}_0 = 0$, and $(\mathcal{M}_t + \mu t)^+$ may be replaced by $\mathcal{M}_t + \mu t$ under the square root in (2.5). Hence, $\mathrm{d}\mathcal{M}_t = \sqrt{2\alpha\beta\mathcal{X}_t}\,\mathrm{d}\mathcal{W}_t$, and the convergences (2.1) and (2.3) can also be formulated as

$$n\left(\widehat{\varrho}_n(\boldsymbol{X}_n)-1\right) \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{X}_t\,\mathrm{d}\mathcal{M}_t - \mathcal{M}_1\int_0^1 \mathcal{X}_t\,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty,$$

$$\widehat{\mu}_n(\boldsymbol{X}_n)-\mu \xrightarrow{\mathcal{L}} \frac{-\int_0^1 \mathcal{X}_t\,\mathrm{d}t\int_0^1 \mathcal{X}_t\,\mathrm{d}\mathcal{M}_t + \mathcal{M}_1\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty.$$

Theorem 2.5. Let $(X_k)_{k\geq -1}$ be an INAR(2) process with autoregressive parameters $(1,0)$ (hence, it is unstable and decomposable). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^4) < \infty$, and $\mu > 0$. Then

$$n^{3/2}\left(\widehat{\varrho}_n(\boldsymbol{X}_n)-1\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0,\ \frac{12\sigma^2}{\mu^2}\right) \qquad \text{as } n\to\infty,$$

$$\begin{bmatrix} n^{1/2}\left(\widehat{\alpha}_n(\boldsymbol{X}_n)-1\right) \\ n^{1/2}\widehat{\beta}_n(\boldsymbol{X}_n) \end{bmatrix} \xrightarrow{\mathcal{L}} Z\begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \text{as } n\to\infty,$$

$$n^{1/2}\left(\widehat{\mu}_n(\boldsymbol{X}_n)-\mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0, \mu^2+4\sigma^2\right) \qquad \text{as } n\to\infty,$$

where Z is a standard normally distributed random variable.

Remark 2.6. Note that an unstable and decomposable INAR(2) process has parameters $(1,0)$; that is, it is actually an unstable INAR(1) process. However, we call the attention that the asymptotic behavior of the estimators $\widehat{\varrho}_n(\boldsymbol{X}_n)$, $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n))$ and $\widehat{\mu}_n(\boldsymbol{X}_n)$ as $n\to\infty$ in Theorem 2.5 can not be derived from the corresponding results for an unstable INAR(1) process, because the CLS estimator of the coefficient (which can also be considered as the stability parameter) of an INAR(1) process is different from $\widehat{\varrho}_n(\boldsymbol{X}_n)$; see, for example, Ispány *et al.* (2003b). Remark that the CLS estimator of the coefficient for an unstable INAR(1) process is also asymptotically normal with the same scaling $n^{3/2}$, but the asymptotic variance $3\sigma^2/\mu^2$ is different from the corresponding one $12\sigma^2/\mu^2$ for an unstable and decomposable INAR(2) process; see Theorem 2.1 of Ispány *et al.* (2003b).

Theorem 2.7. Let $(X_k)_{k\geq -1}$ be an INAR(2) process with autoregressive parameters $(0,1)$ (hence, it is unstable, indecomposable, but not positively regular). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^2) < \infty$, and $\mu > 0$. Then

$$n^{3/2}\left(\widehat{\varrho}_n(\boldsymbol{X}_n)-1\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0,\ \frac{48\sigma^2}{\mu^2}\right) \qquad \text{as } n\to\infty,$$

$$\begin{bmatrix} n\widehat{\alpha}_n(\boldsymbol{X}_n) \\ n(\widehat{\beta}_n(\boldsymbol{X}_n)-1) \end{bmatrix} \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{W}_t\,\mathrm{d}\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2\,\mathrm{d}t}\begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \text{as } n\to\infty,$$

$$n^{1/2}\left(\widehat{\mu}_n(\boldsymbol{X}_n)-\mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0,\ 4\sigma^2\right) \qquad \text{as } n\to\infty,$$

where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process.

Remark 2.8. In each unstable case, the limit distributions of the estimators of (α,β) are concentrated on the same line $\left\{(x,y)\in\mathbb{R}^2 : x+y=0\right\}$. However, these limit distributions are pairwise different. Surprisingly, both in the unstable positively regular case and in the unstable decomposable case, the scaling factor is $\sqrt{n}$, whereas in the unstable, indecomposable, but not positively regular case, it is n. In the stable case, this factor is again $\sqrt{n}$; see Theorem 4.2 of Du

& Li (1991) or Proposition 6.1 of Latour (1998). The reason of this strange phenomena can be understood from the asymptotic behavior of the random sequence $(\boldsymbol{A}_n, \boldsymbol{d}_n)_{n\in\mathbb{N}}$ defined and analyzed in Sections 3–8. Namely, the scaling factors for the entries of the matrices $(\boldsymbol{A}_n)_{n\in\mathbb{N}}$, as well as for the entries of the vectors $(\boldsymbol{d}_n)_{n\in\mathbb{N}}$, are different in the aforementioned cases.

Remark 2.9. The distribution of $\int_0^1 \mathcal{W}_t \, \mathrm{d}\mathcal{W}_t \Big/ \int_0^1 \mathcal{W}_t^2 \, \mathrm{d}t$ in Theorem 2.7 agrees with the limit distribution of the Dickey–Fuller statistics for unit root test of AR(1) time series; see, for example, 17.4.2 and 17.4.7 of Hamilton (1994) or (7.14) and Theorem 9.5.1 of Tanaka (1996). The limit distribution in (2.2) is also a fraction of two stochastic integrals, but it contains two independent standard Wiener processes.

Remark 2.10. We note that the CLS estimators $\widehat{\varrho}_n(\boldsymbol{X}_n)$ and $\left(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n)\right)$ are asymptotically weakly consistent as $n \to \infty$ in Theorems 2.1, 2.5, and 2.7. The CLS estimator $\widehat{\mu}_n(\boldsymbol{X}_n)$ in Theorems 2.5 and 2.7 is also asymptotically weakly consistent as $n \to \infty$; however, in Theorem 2.1, it is not asymptotically weakly consistent. Note that in the case of an unstable INAR(1) model, the CLS estimator of the mean of the innovations is asymptotically weakly consistent; see Ispány *et al.* (2003b). Further, we remark that in Theorem 2.1, the variance σ^2 of the innovations does not show up in the limit distributions, whereas in Theorems 2.5 and 2.7, it appears. Finally, in Theorems 2.1, 2.5, and 2.7, one could prove joint convergence as well.

3. Conditional least squares estimators

Let $\mathcal{F}_k := \sigma(X_{-1}, X_0, \ldots, X_k)$, $k \in \mathbb{Z}_+$. By (1.1),

$$\mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = \alpha X_{k-1} + \beta X_{k-2} + \mu, \qquad k \in \mathbb{N}. \tag{3.1}$$

Let us introduce the sequence

$$M_k := X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}) = X_k - \alpha X_{k-1} - \beta X_{k-2} - \mu, \qquad k \in \mathbb{N}, \tag{3.2}$$

of martingale differences with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$. Then the process $(X_k)_{k\geq -1}$ satisfies the recursion

$$X_k = \alpha X_{k-1} + \beta X_{k-2} + M_k + \mu, \qquad k \in \mathbb{N}. \tag{3.3}$$

For each $n \in \mathbb{N}$, a CLS estimator $\left(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n)\right)$ of the parameters (α, β, μ) based on a sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ can be obtained by minimizing the sum of squares

$$\sum_{k=1}^{n} (X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1}))^2 = \sum_{k=1}^{n} (X_k - \alpha X_{k-1} - \beta X_{k-2} - \mu)^2,$$

with respect to (α, β, μ) over $\mathbb{R}^3$. For all $x_1, \ldots, x_n \in \mathbb{R}$, $n \in \mathbb{N}$, let us put $\boldsymbol{x}_n := (x_1, \ldots, x_n)$, and in what follows, we use the convention $x_{-1} := x_0 := 0$. Consider the function $Q_n : \mathbb{R}^n \times \mathbb{R}^3 \to \mathbb{R}$ given by $Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') := \sum_{k=1}^{n} (x_k - \alpha' x_{k-1} - \beta' x_{k-2} - \mu')^2$ for all $\boldsymbol{x}_n \in \mathbb{R}^n$ and $\alpha', \beta', \mu' \in \mathbb{R}$. A CLS estimator of the parameters (α, β, μ) is a measurable function $\left(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n\right) : \mathbb{R}^n \to \mathbb{R}^3$ such that

$$Q_n\left(\boldsymbol{x}_n; \widehat{\alpha}_n(\boldsymbol{x}_n), \widehat{\beta}_n(\boldsymbol{x}_n), \widehat{\mu}_n(\boldsymbol{x}_n)\right) = \inf_{(\alpha', \beta', \mu')\in\mathbb{R}^3} Q_n\left(\boldsymbol{x}_n; \alpha', \beta', \mu'\right) \qquad \forall\, \boldsymbol{x}_n \in \mathbb{R}^n.$$

Because the variance σ^2 of the innovations does not appear in the conditional expectation $\mathbb{E}(X_k \mid \mathcal{F}_{k-1})$ given in (3.1), and hence, in the definition of Q_n, we do not need to know the value of σ^2 for the calculation of the CLS estimator of the parameters (α, β, μ).

Next, we give the solutions of this extremum problem; for the proof, see Lemma 3.1 of Barczy *et al.* (2012).

Lemma 3.1. *For each $n \in \{2, 3, \ldots\}$, any CLS estimator of the parameters (α, β, μ) is a measurable function $\left(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n\right) : \mathbb{R}^n \to \mathbb{R}^3$ for which*

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_n) \\ \widehat{\mu}_n(\boldsymbol{x}_n) \end{bmatrix} = F_n(\boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_n) \qquad \text{if } \sum_{k=1}^n x_{k-2}^2 > 0,$$

where

$$F_n(\boldsymbol{x}_n) := \sum_{k=1}^n \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix}^\top, \qquad g_n(\boldsymbol{x}_n) := \sum_{k=1}^n x_k \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix},$$

$$\widehat{\alpha}_n(\boldsymbol{x}_n) = \frac{x_n}{x_{n-1}} - \frac{1}{n-1}, \qquad \widehat{\mu}_n(\boldsymbol{x}_n) = \frac{x_{n-1}}{n-1} \qquad \text{if } x_1 = \cdots = x_{n-2} = 0 \text{ and } x_{n-1} \neq 0,$$

and

$$\widehat{\mu}_n(\boldsymbol{x}_n) = \frac{x_n}{n} \qquad \text{if } x_1 = \cdots = x_{n-1} = 0.$$

Note that $\left(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n\right)$ is not defined uniquely on the set $\{\boldsymbol{x}_n \in \mathbb{R}^n : x_1 = \cdots = x_{n-2} = 0\}$. Namely, if $x_1 = \cdots = x_{n-2} = 0$ and $x_{n-1} \neq 0$, then $\widehat{\beta}_n$ can be chosen as an arbitrary measurable function, whereas if $x_1 = \cdots = x_{n-1} = 0$, then the same holds for $\left(\widehat{\alpha}_n, \widehat{\beta}_n\right)$.

Next, we present a result about the existence and uniqueness of $\left(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n)\right)$.

Proposition 3.1. Let $(X_k)_{k \geq -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in [0, 1]^2$ such that $\alpha + \beta = 1$ (hence, it is unstable). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^2) < \infty$, and $\mu > 0$. Then

$$\lim_{n \to \infty} \mathbb{P}\left(\sum_{k=1}^n X_{k-2}^2 > 0\right) = 1,$$

and thus the probability of the existence of a unique CLS estimator $\left(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n)\right)$ converges to 1 as $n \to \infty$, and this CLS estimator has the form

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \boldsymbol{F}_n^{-1} \boldsymbol{g}_n,$$

on the set $\left\{\omega \in \Omega : \sum_{k=1}^n X_{k-2}(\omega)^2 > 0\right\}$, where

$$\boldsymbol{F}_n := \sum_{k=1}^n \begin{bmatrix} X_{k-1}^2 & X_{k-1}X_{k-2} & X_{k-1} \\ X_{k-1}X_{k-2} & X_{k-2}^2 & X_{k-2} \\ X_{k-1} & X_{k-2} & 1 \end{bmatrix}, \qquad \boldsymbol{g}_n := \sum_{k=1}^n \begin{bmatrix} X_k X_{k-1} \\ X_k X_{k-2} \\ X_k \end{bmatrix}.$$

Proof. First, we prove the statements for $(\alpha, \beta) \in (0,1)^2$. For each $n \in \mathbb{N}$, consider the random step process $\mathcal{X}_t^{(n)} := n^{-1} X_{\lfloor nt \rfloor}$, $t \in \mathbb{R}_+$. By Theorem 3.1 of Barczy *et al.* (2011), we have

$$\mathcal{X}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{X} \qquad \text{as } n \to \infty, \tag{3.4}$$

where the process $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ is the unique strong solution of the SDE (2.4) with initial value $\mathcal{X}_0 = 0$. By (3.4), Lemmas B.2, and B.3, one can show

$$\frac{1}{n^3} \sum_{k=1}^{n} X_{k-2}^2 \xrightarrow{\mathcal{L}} \int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t \qquad \text{as } n \to \infty; \tag{3.5}$$

see Proposition 3.1 of Barczy *et al.* (2012). Because $\mu > 0$, by the SDE (2.4), we have $\mathbb{P}(\mathcal{X}_t = 0,\ t \in [0,1]) = 0$, which implies that $\mathbb{P}\left(\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t > 0\right) = 1$. Consequently, the distribution function of $\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t$ is continuous at 0, and hence, by (3.5),

$$\mathbb{P}\left(\sum_{k=1}^{n} X_{k-2}^2 > 0\right) = \mathbb{P}\left(\frac{1}{n^3}\sum_{k=1}^{n} X_{k-2}^2 > 0\right) \to \mathbb{P}\left(\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t > 0\right) = 1$$

as $n \to \infty$, which implies the statement in the case of $(\alpha, \beta) \in (0,1)^2$.

Next, we consider the case of $(\alpha, \beta) = (1, 0)$. In this case, (1.1) has the form $X_n = X_{n-1} + \varepsilon_n$, $n \in \mathbb{N}$, and hence, $X_n = \sum_{k=1}^{n} \varepsilon_k$, $n \in \mathbb{N}$. By the strong law of large numbers, we have

$$n^{-1} X_n \xrightarrow{\text{a.s.}} \mu,$$

and, hence $n^{-2} X_n^2 \xrightarrow{\text{a.s.}} \mu^2$, where $\xrightarrow{\text{a.s.}}$ denotes almost sure convergence. Then, by the Toeplitz theorem, we conclude

$$n^{-3} \sum_{k=1}^{n} X_k^2 \xrightarrow{\text{a.s.}} \frac{1}{3}\mu^2.$$

Because $\mu > 0$, this implies the existence of an event $\Omega_0 \in \mathcal{A}$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$, there exists $n_0(\omega) \in \mathbb{N}$ such that $\sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0$ for $n \geq n_0(\omega)$. This is equivalent with $1 = \mathbb{P}(\bigcup_{n=1}^{\infty}\{\sum_{k=1}^{n} X_{k-2}^2 > 0\}) = \lim_{n\to\infty} \mathbb{P}(\{\sum_{k=1}^{n} X_{k-2}^2 > 0\})$; hence, we obtain the statement in case $(\alpha, \beta) = (1, 0)$.

Finally, we consider the case of $(\alpha, \beta) = (0, 1)$. In this case, (1.1) has the form $X_k = X_{k-2} + \varepsilon_k$, $k \in \mathbb{N}$, and hence, $X_{2n} = \sum_{k=1}^{n} \varepsilon_{2k}$, $X_{2n-1} = \sum_{k=1}^{n} \varepsilon_{2k-1}$, $n \in \mathbb{N}$. By the strong law of large numbers, we have $n^{-1} X_{2n} \xrightarrow{\text{a.s.}} \mu$ and $n^{-1} X_{2n-1} \xrightarrow{\text{a.s.}} \mu$, which yield that

$$n^{-1} X_n \xrightarrow{\text{a.s.}} \frac{\mu}{2}. \tag{3.6}$$

Using the Toeplitz theorem, as in case $(\alpha, \beta) = (1, 0)$, we obtain

$$n^{-3} \sum_{k=1}^{n} X_k^2 \xrightarrow{\text{a.s.}} \frac{1}{12}\mu^2. \tag{3.7}$$

One can finish the proof as in case $(\alpha, \beta) = (1, 0)$. □

The recursion (3.3) can also be written in the form $X_k = \varrho X_{k-1} - \beta V_{k-1} + M_k + \mu$, $k \in \mathbb{N}$, with

$$V_{k-1} := X_{k-1} - X_{k-2}, \qquad k \in \mathbb{N}.$$

This reparametrization can be called the canonical form of Sims *et al.* (1990); see also 17.7.6 in Hamilton (1994). One can check again, see Barczy *et al.* (2012), that for an unstable INAR(2) process, that is, when $\varrho = 1$, the probability of the existence of a unique CLS estimator $(\widehat{\varrho}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ converges to 1 as $n \to \infty$; this CLS estimator has the form

$$\begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \boldsymbol{A}_n^{-1}\boldsymbol{b}_n, \tag{3.8}$$

on the set $\{\omega \in \Omega : \sum_{k=1}^n X_{k-2}(\omega)^2 > 0\}$, where

$$\boldsymbol{A}_n := \sum_{k=1}^n \begin{bmatrix} X_{k-1}^2 & -X_{k-1}V_{k-1} & X_{k-1} \\ -X_{k-1}V_{k-1} & V_{k-1}^2 & -V_{k-1} \\ X_{k-1} & -V_{k-1} & 1 \end{bmatrix}, \qquad \boldsymbol{b}_n := \sum_{k=1}^n \begin{bmatrix} X_k X_{k-1} \\ -X_k V_{k-1} \\ X_k \end{bmatrix},$$

and $\widehat{\varrho}_n(\boldsymbol{X}_n) = \widehat{\alpha}_n(\boldsymbol{X}_n) + \widehat{\beta}_n(\boldsymbol{X}_n)$. Note also that in case of an unstable INAR(2) process, that is, when $\varrho = 1$, we have

$$V_k = -\beta V_{k-1} + M_k + \mu, \qquad k \in \mathbb{N}, \tag{3.9}$$

hence, $(V_k)_{k\in\mathbb{Z}_+}$ is a stable AR(1) process with heteroscedastic innovations $(M_k)_{k\in\mathbb{N}}$ and with positive drift μ whenever $0 \le \beta < 1$.

4. Proof of the main results

In case of an unstable INAR(2) process, that is, when $\varrho = \alpha + \beta = 1$, by (3.8), we have

$$\begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\beta}_n(\boldsymbol{X}_n) - \beta \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} = \boldsymbol{A}_n^{-1}\boldsymbol{d}_n, \qquad n \in \mathbb{N},$$

on the set $\{\omega \in \Omega : \sum_{k=1}^n X_{k-2}(\omega)^2 > 0\}$, where

$$\boldsymbol{d}_n := \sum_{k=1}^n \begin{bmatrix} M_k X_{k-1} \\ -M_k V_{k-1} \\ M_k \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Theorems 2.1, 2.5, and 2.7 will follow by the continuous mapping theorem from Theorems 4.1, 4.2, and 4.3, respectively; see the details in Barczy *et al.* (2012).

Theorem 4.1. Under the assumptions of Theorem 2.1, $(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \xrightarrow{\mathcal{L}} (\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}})$ as $n \to \infty$, where

$$\widetilde{\boldsymbol{A}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{A}_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \quad \widetilde{\boldsymbol{d}}_n := \begin{bmatrix} n^{-2} & 0 & 0 \\ 0 & n^{-3/2} & 0 \\ 0 & 0 & n^{-1} \end{bmatrix} \boldsymbol{d}_n,$$

$$\widetilde{\boldsymbol{A}} := \begin{bmatrix} \int_0^1 \mathcal{X}_t^2\,\mathrm{d}t & 0 & \int_0^1 \mathcal{X}_t\,\mathrm{d}t \\ 0 & \frac{2\beta}{1+\beta}\int_0^1 \mathcal{X}_t\,\mathrm{d}t & 0 \\ \int_0^1 \mathcal{X}_t\,\mathrm{d}t & 0 & 1 \end{bmatrix}, \qquad \widetilde{\boldsymbol{d}} := \begin{bmatrix} \sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t^{3/2}\,\mathrm{d}\mathcal{W}_t \\ -\frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\int_0^1 \mathcal{X}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t \\ (1+\beta)\mathcal{X}_1 - \mu \end{bmatrix},$$

and $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes.

Theorem 4.2. Under the assumptions of Theorem 2.5, $(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \xrightarrow{\mathcal{L}} (\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}})$ as $n \to \infty$, where

$$\widetilde{\boldsymbol{A}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{A}_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \quad \widetilde{\boldsymbol{d}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{d}_n,$$

$$\widetilde{\boldsymbol{A}} := \begin{bmatrix} \frac{1}{3}\mu^2 & -\frac{1}{2}\mu^2 & \frac{1}{2}\mu \\ -\frac{1}{2}\mu^2 & \mu^2+\sigma^2 & -\mu \\ \frac{1}{2}\mu & -\mu & 1 \end{bmatrix}, \qquad \widetilde{\boldsymbol{d}} \stackrel{\mathcal{L}}{=} \mathcal{N}_3(\boldsymbol{0}, \sigma^2 \widetilde{\boldsymbol{A}}),$$

where $\stackrel{\mathcal{L}}{=}$ means equality in distribution.

Theorem 4.3. Under the assumptions of Theorem 2.7, $(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \xrightarrow{\mathcal{L}} (\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}})$ as $n \to \infty$, where

$$\widetilde{\boldsymbol{A}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{A}_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \widetilde{\boldsymbol{d}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{d}_n,$$

$$\widetilde{\boldsymbol{A}} := \begin{bmatrix} \frac{1}{12}\mu^2 & 0 & \frac{1}{4}\mu \\ 0 & \sigma^2 \int_0^1 \mathcal{W}_t^2 \,\mathrm{d}t & 0 \\ \frac{1}{4}\mu & 0 & 1 \end{bmatrix}, \qquad \widetilde{\boldsymbol{d}} := \begin{bmatrix} \frac{1}{2}\mu\sigma \int_0^1 t \,\mathrm{d}\widetilde{\mathcal{W}}_t \\ \sigma^2 \int_0^1 \mathcal{W}_t \,\mathrm{d}\mathcal{W}_t \\ \sigma \widetilde{\mathcal{W}}_1 \end{bmatrix},$$

and $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes.

5. Proof of Theorem 4.1

We have

$$\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-3} X_{k-1}^2 & -n^{-5/2} X_{k-1} V_{k-1} & n^{-2} X_{k-1} \\ -n^{-5/2} X_{k-1} V_{k-1} & n^{-2} V_{k-1}^2 & -n^{-3/2} V_{k-1} \\ n^{-2} X_{k-1} & -n^{-3/2} V_{k-1} & n^{-1} \end{bmatrix}, \tag{5.1}$$

$$\widetilde{\boldsymbol{d}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-2} M_k X_{k-1} \\ -n^{-3/2} M_k V_{k-1} \\ n^{-1} M_k \end{bmatrix}. \tag{5.2}$$

Lemma 5.1. *Under the assumptions of Theorem 2.1, we have*

$$n^{-3/2} \sum_{k=1}^{n} V_k \xrightarrow{\mathbb{P}} 0, \quad n^{-5/2} \sum_{k=1}^{n} X_k V_k \xrightarrow{\mathbb{P}} 0, \quad n^{-2} \left(\sum_{k=1}^{n} V_k^2 - \frac{2\beta}{1+\beta} \sum_{k=1}^{n} X_{k-1} \right) \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$, where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability.

Proof. We have $\sum_{k=1}^{n} V_k = X_n \geq 0$ and, by Corollary A.4, $\mathbb{E}(X_n) = \mathrm{O}(n)$; hence, we conclude the first convergence. Corollary A.4 implies the second convergence, because

$$\sum_{k=1}^{n} X_k V_k = \frac{1}{2} X_n^2 + \frac{1}{2} \sum_{k=1}^{n} V_k^2 \geq 0,$$

$$\mathbb{E}\left(\sum_{k=1}^{n} X_k V_k \right) = \frac{1}{2}\mathbb{E}(X_n^2) + \frac{1}{2} \sum_{k=1}^{n} \mathbb{E}(V_k^2) = \mathrm{O}(n^2).$$

In order to prove the last statement, we derive a decomposition of $\sum_{k=1}^{n} V_k^2$ as a sum of a martingale and some negligible terms. Using recursion (3.9) and Lemma A.1, we obtain

$$\begin{aligned}\mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) &= \beta^2 V_{k-1}^2 + 2\alpha\beta X_{k-1} + \mu^2 + \sigma^2 - (2\beta\mu + \alpha\beta)V_{k-1}, \\ &= \beta^2 V_{k-1}^2 + 2\alpha\beta X_{k-1} + \text{constant} + \text{constant} \times V_{k-1}.\end{aligned}$$

Thus,

$$\sum_{k=1}^{n} V_k^2 = \sum_{k=1}^{n} \Big[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\Big] + \beta^2 \sum_{k=1}^{n} V_{k-1}^2 + 2\alpha\beta \sum_{k=1}^{n} X_{k-1} + \mathrm{O}(n) + \text{const.} \times \sum_{k=1}^{n} V_{k-1}.$$

Consequently,

$$\begin{aligned}\sum_{k=1}^{n} V_k^2 = \frac{1}{1-\beta^2} \sum_{k=1}^{n} \Big[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\Big] + \frac{2\beta}{1+\beta} \sum_{k=1}^{n} X_{k-1} \\ - \frac{\beta^2}{1-\beta^2} V_n^2 + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{n} V_{k-1}.\end{aligned} \tag{5.3}$$

By (A.7) with $(\ell, i, j) = (8, 0, 2)$, we obtain $n^{-2} \sum_{k=1}^{n} \big[V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})\big] \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$. By Corollary A.4, $\mathbb{E}(V_n^2) = \mathrm{O}(n)$ and $\mathbb{E}(X_{n-1}^2) = \mathrm{O}(n^2)$, and because $\sum_{k=1}^{n} V_{k-1} = X_{n-1}$, we obtain $n^{-2} V_n^2 \xrightarrow{\mathbb{P}} 0$ and $n^{-2} \sum_{k=1}^{n} V_{k-1} \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$. Hence, by (5.3), we obtain the last statement. □

Now let

$$U_k := X_k + \beta X_{k-1}, \qquad k \in \mathbb{Z}_+,$$

with the convention $U_{-1} := U_0 := 0$. One can observe that $U_k \geq 0$ for all $k \in \mathbb{Z}_+$, and by $\alpha + \beta = 1$, $U_k = U_{k-1} + M_k + \mu$, $k \in \mathbb{Z}_+$; hence, $(U_k)_{k \in \mathbb{Z}_+}$ is a non-negative unstable AR(1) process with positive drift μ sharing the innovations $(M_k)_{k \in \mathbb{N}}$ with the stable AR(1) process $(V_k)_{k \in \mathbb{Z}_+}$.

Consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \mathcal{M}_t^{(n)} \\ \mathcal{N}_t^{(n)} \\ \mathcal{P}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Z}_k^{(n)} \qquad \text{with} \qquad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1} M_k \\ n^{-2} M_k U_{k-1} \\ n^{-3/2} M_k V_{k-1} \end{bmatrix},$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$. Theorem 4.1 will follow from Lemma 5.1 and the following theorem (this will be detailed after Theorem 5.1).

Theorem 5.1. Under the assumptions of Theorem 2.1, we have

$$\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{Z}} \qquad \text{as } n \to \infty, \tag{5.4}$$

where the process $(\boldsymbol{\mathcal{Z}}_t)_{t \in \mathbb{R}_+}$ with values in $\mathbb{R}^3$ is the unique strong solution of the SDE

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t)\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \tag{5.5}$$

with $\boldsymbol{\mathcal{Z}}_0 = \mathbf{0}$, where $\boldsymbol{\mathcal{W}}_t := \begin{bmatrix} \mathcal{W}_t & \widetilde{\mathcal{W}}_t \end{bmatrix}^\top$, $t \in \mathbb{R}_+$, is a two-dimensional standard Wiener process, and $\gamma : \mathbb{R}_+ \times \mathbb{R}^3 \to \mathbb{R}^{3\times 2}$ is given by

$$\gamma(t, \boldsymbol{x}) := \begin{bmatrix} \sqrt{\frac{2\alpha\beta}{1+\beta}}[(x_1 + \mu t)^+]^{1/2} & 0 \\ \sqrt{\frac{2\alpha\beta}{1+\beta}}[(x_1 + \mu t)^+]^{3/2} & 0 \\ 0 & \frac{2\beta\sqrt{\alpha}}{(1+\beta)^{3/2}}(x_1 + \mu t) \end{bmatrix},$$

for $t \in \mathbb{R}_+$ and $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Indeed, the unique strong solution of (5.5) with initial value $\boldsymbol{\mathcal{Z}}_0 = \mathbf{0}$ is

$$\boldsymbol{\mathcal{Z}}_t =: \begin{bmatrix} \mathcal{M}_t \\ \mathcal{N}_t \\ \mathcal{P}_t \end{bmatrix} = \begin{bmatrix} (1+\beta)\mathcal{X}_t - \mu t \\ (1+\beta)\sqrt{2\alpha\beta}\int_0^t \mathcal{X}_s^{3/2}\,\mathrm{d}\mathcal{W}_s \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\int_0^t \mathcal{X}_s\,\mathrm{d}\widetilde{\mathcal{W}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+,$$

because by Remark 2.4, (5.5) can be written in the form

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \sqrt{\frac{2\alpha\beta}{1+\beta}}[(\mathcal{M}_t + \mu t)^+]^{1/2}\,\mathrm{d}\mathcal{W}_t \\ \sqrt{\frac{2\alpha\beta}{1+\beta}}[(\mathcal{M}_t + \mu t)^+]^{3/2}\,\mathrm{d}\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{(1+\beta)^{3/2}}(\mathcal{M}_t + \mu t)\,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix} = \begin{bmatrix} \sqrt{2\alpha\beta}\mathcal{X}_t^{1/2}\,\mathrm{d}\mathcal{W}_t \\ (1+\beta)\sqrt{2\alpha\beta}\mathcal{X}_t^{3/2}\,\mathrm{d}\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\mathcal{X}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}.$$

By the method of the proof of $\mathcal{X}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{X}$ in Theorem 3.1 in Barczy *et al.* (2011), using a functional version of Lemma B.2, one can easily derive

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \mathcal{X} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty; \tag{5.6}$$

see page 19 in Barczy *et al.* (2012). Next, similarly to the proof of (3.5), by Lemmas B.2 and B.3, convergence (5.6) implies

$$\sum_{k=1}^n \begin{bmatrix} n^{-1}M_k \\ n^{-3}X_{k-1}^2 \\ n^{-2}X_{k-1} \\ n^{-2}M_kU_{k-1} \\ n^{-3/2}M_kV_{k-1} \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} (1+\beta)\mathcal{X}_1 - \mu \\ \int_0^1 \mathcal{X}_t^2\,\mathrm{d}t \\ \int_0^1 \mathcal{X}_t\,\mathrm{d}t \\ (1+\beta)\sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t^{3/2}\,\mathrm{d}\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\int_0^1 \mathcal{X}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix} \qquad \text{as } n \to \infty.$$

Using $U_{k-1} = (1+\beta)X_{k-1} - \beta V_{k-1}$ and convergence of the third coordinates in $\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{Z}}$, we obtain

$$n^{-2}\left(\sum_{k=1}^n M_kX_{k-1} - \frac{1}{1+\beta}\sum_{k=1}^n M_kU_{k-1}\right) = \frac{\beta}{(1+\beta)n^2}\sum_{k=1}^n M_kV_{k-1} \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$. Using (5.1), (5.2), the aforementioned two convergences, and Lemma 5.1, we obtain Theorem 4.1 by Slutsky's lemma.

6. Proof of Theorem 5.1

In order to show convergence $\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{Z}}$, we apply Theorem C.1. Note that the arguments in Section 5 and Remark 2.3 show that the SDE (5.5) admits a unique strong solution $(\boldsymbol{\mathcal{Z}}_t^{z})_{t\in\mathbb{R}_+}$ for all initial values $\boldsymbol{\mathcal{Z}}_0^{z} = z \in \mathbb{R}^3$.

Now, we show that conditions (i) and (ii) of Theorem C.1 hold. The conditional variance matrix $\mathbb{E}\left(\boldsymbol{Z}_k^{(n)}(\boldsymbol{Z}_k^{(n)})^\top \mid \mathcal{F}_{k-1}\right)$ has the form

$$\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) \begin{bmatrix} n^{-2} & n^{-3}U_{k-1} & n^{-5/2}V_{k-1} \\ n^{-3}U_{k-1} & n^{-4}U_{k-1}^2 & n^{-7/2}U_{k-1}V_{k-1} \\ n^{-5/2}V_{k-1} & n^{-7/2}U_{k-1}V_{k-1} & n^{-3}V_{k-1}^2 \end{bmatrix},$$

for $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, and the matrix $\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})^\top$ equals

$$\begin{bmatrix} \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s) & \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^2 & 0 \\ \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^2 & \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^3 & 0 \\ 0 & 0 & \frac{4\alpha\beta^2}{(1+\beta)^3}(\mathcal{M}_s^{(n)} + \mu s)^2 \end{bmatrix},$$

for $s \in \mathbb{R}_+$, where we used that $(\mathcal{M}_s^{(n)} + \mu s)^+ = \mathcal{M}_s^{(n)} + \mu s$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$; see page 598 in Barczy *et al.* (2011) or (6.7) later on. In order to check condition (i) of Theorem C.1, we need to prove that, for each $T > 0$,

$$\sup_{t\in[0,T]} \left| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)\,\mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0, \tag{6.1}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2\,\mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0, \tag{6.2}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^3\,\mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0, \tag{6.3}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{4\alpha\beta^2}{(1+\beta)^3} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2\,\mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0, \tag{6.4}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) \right| \xrightarrow{\mathbb{P}} 0, \tag{6.5}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1}\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) \right| \xrightarrow{\mathbb{P}} 0 \tag{6.6}$$

as $n \to \infty$. Covergence (6.1) follows from (5.1) in Barczy *et al.* (2011) with the special choices $p = 2$, $\alpha_1 = \alpha$ and $\alpha_2 = \beta$.

Next, we turn to prove (6.2). We have

$$\mathcal{M}_s^{(n)} + \mu s = \frac{1}{n}U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n}\mu, \qquad s \in \mathbb{R}_+, \quad n \in \mathbb{N}. \tag{6.7}$$

Thus,

$$\int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2 \, \mathrm{d}s = \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k^2 + \frac{\mu}{n^3} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k + \frac{nt - \lfloor nt \rfloor}{n^3} U_{\lfloor nt \rfloor}^2 + \frac{\mu (nt - \lfloor nt \rfloor)^2}{n^3} U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^3}{3n^3} \mu^2.$$

Because

$$X_{k-1} = \frac{1}{1+\beta}(U_k - V_k), \qquad X_k = \frac{1}{1+\beta}(U_k + \beta V_k), \qquad k \in \mathbb{N}, \tag{6.8}$$

using Lemma A.1, we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) = \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 - \frac{\alpha^2 \beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} + \sigma^2 \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}. \tag{6.9}$$

Thus, in order to show (6.2), it suffices to prove that

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0, \tag{6.10}$$

$$n^{-3/2} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sup_{t \in [0,T]} \left[\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^3 \right] \to 0 \tag{6.11}$$

as $n \to \infty$. Using (A.5) with $(\ell, i, j) = (8, 1, 1)$ and $(\ell, i, j) = (8, 1, 0)$, we obtain (6.10). Using (A.6) with $(\ell, i, j) = (8, 1, 0)$, and $|nt - \lfloor nt \rfloor| \le 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, we conclude (6.11). Convergence (6.3) can be checked in a similar way.

Next, we turn to prove (6.4). By (6.9) and (6.10), we obtain

$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$ for all $T > 0$. Using (6.2), in order to prove (6.4), it is sufficient to show that

$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{4\alpha\beta^2}{(1+\beta)^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0, \tag{6.12}$$

as $n \to \infty$ for all $T > 0$. As in the previous case, using Lemma A.1 and (6.8), we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) = \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{\alpha^2 \beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^3 + \sigma^2 \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2.$$

Using (A.5) with $(\ell, i, j) = (8, 0, 3)$ and $(\ell, i, j) = (8, 0, 2)$, we have

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k|^3 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty;$$

hence, (6.12) will follow from

$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{2\beta}{(1+\beta)^2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0, \tag{6.13}$$

as $n \to \infty$ for all $T > 0$.

In what follows, we decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - 2\beta(1+\beta)^{-2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2$ as a sum of a martingale and some negligible terms. Using the method of the proof of Lemma 5.1, we obtain

$$\begin{aligned}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 = \frac{1}{1-\beta^2} \sum_{k=2}^{\lfloor nt \rfloor} \Big[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \mid \mathcal{F}_{k-2}) \Big] \\ + \frac{2\beta}{(1+\beta)^2} \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2 - \frac{\beta^2}{1-\beta^2} U_{\lfloor nt \rfloor - 1} V_{\lfloor nt \rfloor - 1}^2 + \mathrm{O}(n) \\ + \text{lin. comb. of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}, \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}^2, \ \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.\end{aligned}$$

Using (A.7) with $(\ell, i, j) = (8, 1, 2)$, we have

$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=2}^{\lfloor nt \rfloor} \Big[U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \mid \mathcal{F}_{k-2}) \Big] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (6.13), it suffices to prove that

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0, \tag{6.14}$$

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0, \tag{6.15}$$

$$n^{-3} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} V_{\lfloor nt \rfloor}^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor}^2 \xrightarrow{\mathbb{P}} 0 \tag{6.16}$$

as $n \to \infty$. Using (A.5) with $(\ell, i, j) = (8, 1, 1)$, $(\ell, i, j) = (8, 0, 2)$, $(\ell, i, j) = (8, 1, 0)$, and $(\ell, i, j) = (8, 0, 1)$, we obtain (6.14) and (6.15). Using (A.6) with $(\ell, i, j) = (8, 1, 2)$ and $(\ell, i, j) = (8, 2, 0)$, we have (6.16). Thus, we conclude (6.4). Covergences (6.5) and (6.6) can be proved similarly.

Finally, we check condition (ii) of Theorem C.1, that is, the conditional Lindeberg condition, namely, for all $\theta > 0$ and $T > 0$,

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \ {}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1} \right) \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty. \tag{6.17}$$

This will follow from $\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \ {}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \right) \to 0$ as $n \to \infty$. We have

$$\mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \ {}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \right) \le \theta^{-2} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^4 \right) \le \theta^{-2} \mathbb{E}\left(3 M_k^4 \left(n^{-4} + n^{-8} U_{k-1}^4 + n^{-6} V_{k-1}^4 \right) \right).$$

By Corollary A.4, we have $\mathbb{E}(M_k^4) = O(k^2)$, $\mathbb{E}(M_k^4 U_{k-1}^4) \le \sqrt{\mathbb{E}(M_k^8)\mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)$ and $\mathbb{E}(M_k^4 V_{k-1}^4) \le \sqrt{\mathbb{E}(M_k^8)\mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)$; hence, we obtain (6.17).

7. Proof of Theorem 4.2

We have

$$\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-3}X_{k-1}^2 & -n^{-2}X_{k-1}V_{k-1} & n^{-2}X_{k-1} \\ -n^{-2}X_{k-1}V_{k-1} & n^{-1}V_{k-1}^2 & -n^{-1}V_{k-1} \\ n^{-2}X_{k-1} & -n^{-1}V_{k-1} & n^{-1} \end{bmatrix},$$

$$\widetilde{\boldsymbol{d}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-3/2}M_k X_{k-1} \\ -n^{-1/2}M_k V_{k-1} \\ n^{-1/2}M_k \end{bmatrix}.$$

Theorem 4.2 will follow from the following statement and Slutsky's lemma.

Theorem 7.1. Under the assumptions of Theorem 2.5, we have

$$n^{-2}\sum_{k=1}^{n} X_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{2}, \qquad n^{-1}\sum_{k=1}^{n} V_{k-1} \xrightarrow{\text{a.s.}} \mu, \qquad n^{-3}\sum_{k=1}^{n} X_{k-1}^2 \xrightarrow{\text{a.s.}} \frac{\mu^2}{3},$$

$$n^{-2}\sum_{k=1}^{n} X_{k-1}V_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu^2}{2}, \qquad n^{-1}\sum_{k=1}^{n} V_{k-1}^2 \xrightarrow{\text{a.s.}} \sigma^2+\mu^2,$$

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3/2}M_k X_{k-1} \\ -n^{-1/2}M_k V_{k-1} \\ n^{-1/2}M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_3\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} \frac{1}{3}\mu^2 & -\frac{1}{2}\mu^2 & \frac{1}{2}\mu \\ -\frac{1}{2}\mu^2 & \sigma^2+\mu^2 & -\mu \\ \frac{1}{2}\mu & -\mu & 1 \end{bmatrix} \right).$$

Proof. In this case, (1.1) has the form $X_k = X_{k-1} + \varepsilon_k$, $k \in \mathbb{N}$, and hence, $X_k = \varepsilon_1 + \cdots + \varepsilon_k$, $M_k = X_k - X_{k-1} - \mu = \varepsilon_k - \mu$, and $V_k = X_k - X_{k-1} = \varepsilon_k$, $k \in \mathbb{N}$. The statements are easy applications of the strong law of large numbers, the Toeplitz theorem, and the martingale central limit theorem; see Theorem 7.1 in Barczy *et al.* (2012). □

8. Proof of Theorem 4.3

We have

$$\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-3}X_{k-1}^2 & -n^{-5/2}X_{k-1}V_{k-1} & n^{-2}X_{k-1} \\ -n^{-5/2}X_{k-1}V_{k-1} & n^{-2}V_{k-1}^2 & -n^{-3/2}V_{k-1} \\ n^{-2}X_{k-1} & -n^{-3/2}V_{k-1} & n^{-1} \end{bmatrix},$$

$$\widetilde{\boldsymbol{d}}_n = \sum_{k=1}^{n} \begin{bmatrix} n^{-3/2}M_k X_{k-1} \\ -n^{-1}M_k V_{k-1} \\ n^{-1/2}M_k \end{bmatrix}.$$

Lemma 8.1. *Under the assumptions of Theorem 2.7, as $n \to \infty$, we have*

$$n^{-2}\sum_{k=1}^{n} X_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{4}, \qquad n^{-1}\sum_{k=1}^{n} V_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{2}, \qquad n^{-3}\sum_{k=1}^{n} X_{k-1}^2 \xrightarrow{\text{a.s.}} \frac{\mu^2}{12},$$

$$n^{-5/2}\sum_{k=1}^{n} X_{k-1}V_{k-1} \xrightarrow{\mathbb{P}} 0, \qquad n^{-2}\sum_{k=1}^{n} [\mathbb{E}(V_{k-1})]^2 \to 0, \qquad n^{-1}\sum_{k=1}^{n} M_k \mathbb{E}(V_{k-1}) \xrightarrow{\mathbb{P}} 0,$$

$$n^{-2}\sum_{k=1}^{n} (V_{k-1} - \mathbb{E}(V_{k-1}))\mathbb{E}(V_{k-1}) \xrightarrow{\mathbb{P}} 0, \qquad n^{-3/2}\sum_{k=1}^{n} M_k\,(X_{k-1} - \mathbb{E}(X_{k-1})) \xrightarrow{\mathbb{P}} 0.$$

Proof. In this case, (1.1) has the form $X_k = X_{k-2} + \varepsilon_k$, $k \in \mathbb{N}$, and hence, $X_{2k} = \varepsilon_2 + \varepsilon_4 + \cdots + \varepsilon_{2k}$, $X_{2k-1} = \varepsilon_1 + \varepsilon_3 + \cdots + \varepsilon_{2k-1}$, $M_k = X_k - X_{k-2} - \mu = \varepsilon_k - \mu$, and $V_{2k} = X_{2k} - X_{2k-1} = (\varepsilon_2 - \varepsilon_1) + \cdots + (\varepsilon_{2k} - \varepsilon_{2k-1})$, $V_{2k-1} = X_{2k-1} - X_{2k-2} = (\varepsilon_1 - \varepsilon_2) + \cdots + (\varepsilon_{2k-3} - \varepsilon_{2k-2}) + \varepsilon_{2k-1}$, $k \in \mathbb{N}$. The first convergence follows from (3.6) by the Toeplitz theorem. Again by (3.6), we obtain $n^{-1} \sum_{k=1}^{n} V_{k-1} = n^{-1} X_{n-1} \xrightarrow{\text{a.s.}} \mu/2$. We have already have shown the third convergence; see (3.7). The fourth convergence can be obtained similarly as the second convergence in Lemma 5.1. For each $k \in \mathbb{N}$, we have $\mathbb{E}(V_{2k}) = 0$ and $\mathbb{E}(V_{2k-1}) = \mu$; hence, we conclude the fifth convergence. Because $\mathbb{E}(V_{2k}) = 0$ and $\mathbb{E}(V_{2k-1}) = \mu$, $k \in \mathbb{N}$, we have

$$\mathbb{E}\left(\left(\sum_{k=1}^{n} M_k \mathbb{E}(V_{k-1})\right)^2\right) = \sigma^2 \sum_{k=1}^{n} [\mathbb{E}(V_{k-1})]^2 = \mathrm{O}(n),$$

which implies the sixth convergence. Further,

$$\mathbb{E}\left(\left|\sum_{k=1}^{n} (V_{k-1} - \mathbb{E}(V_{k-1}))\mathbb{E}(V_{k-1})\right|\right) \leq \sum_{k=1}^{n} \mu \sqrt{\mathbb{E}\left((V_{k-1} - \mathbb{E}(V_{k-1}))^2\right)},$$

which is of order $\mathrm{O}(n^{3/2})$, because $\mathbb{E}(V_k - \mathbb{E}(V_k))^2 \leq \mathbb{E}(V_k^2) = \mathrm{O}(k)$, $k \in \mathbb{N}$, by Corollary A.4; thus, we obtain the seventh convergence. Moreover, using that $M_k(X_{k-1} - \mathbb{E}(X_{k-1})) = (\varepsilon_k - \mu)(X_{k-1} - \mathbb{E}(X_{k-1}))$, $k \in \{1, \ldots, n\}$, are uncorrelated,

$$\mathbb{E}\left(\left(\sum_{k=1}^{n} M_k\, (X_{k-1} - \mathbb{E}(X_{k-1}))\right)^2\right) = \sigma^2 \sum_{k=1}^{n} \mathbb{E}\left((X_{k-1} - \mathbb{E}(X_{k-1}))^2\right),$$

which is of order $\mathrm{O}(n^2)$, because $\mathbb{E}(X_{k-1} - \mathbb{E}(X_{k-1}))^2 = \lfloor k/2 \rfloor \sigma^2$, $k \in \mathbb{N}$; thus, we obtain the last convergence. □

Theorem 4.3 will follow from Lemma 8.1 and the following statement.

Theorem 8.1. Under the assumptions of Theorem 2.7, we have

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-2}\,(V_{k-1} - \mathbb{E}(V_{k-1}))^2 \\ n^{-3/2} M_k \mathbb{E}(X_{k-1}) \\ -n^{-1} M_k\,(V_{k-1} - \mathbb{E}(V_{k-1})) \\ n^{-1/2} M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \sigma^2 \int_0^1 \mathcal{W}_t^2\, \mathrm{d}t \\ \frac{1}{2}\mu\sigma \int_0^1 t\, \mathrm{d}\widetilde{\mathcal{W}}_t \\ \sigma^2 \int_0^1 \mathcal{W}_t\, \mathrm{d}\mathcal{W}_t \\ \sigma \widetilde{\mathcal{W}}_1 \end{bmatrix} \qquad \text{as } n \to \infty,$$

where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes.

Proof. Consider the sequence

$$\begin{bmatrix} \mathcal{S}_t^{(n)} \\ \mathcal{T}_t^{(n)} \end{bmatrix} := \begin{bmatrix} n^{-1/2}\left(X_{2\lfloor nt \rfloor} - \mathbb{E}(X_{2\lfloor nt \rfloor})\right) \\ n^{-1/2}\left(X_{2\lfloor nt \rfloor - 1} - \mathbb{E}(X_{2\lfloor nt \rfloor - 1})\right) \end{bmatrix}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N},$$

of stochastic processes. Then, by the martingale central limit theorem,

$$\begin{bmatrix} \mathcal{S}^{(n)} \\ \mathcal{T}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{L}} \sigma \begin{bmatrix} \mathcal{B} \\ \widetilde{\mathcal{B}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{8.1}$$

where $(\mathcal{B}_t)_{t\in\mathbb{R}_+}$ and $(\widetilde{\mathcal{B}}_t)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes.

Next, we are going to prove that convergence (8.1) implies

$$\sum_{k=1}^{n}\begin{bmatrix} n^{-2}\,(V_{k-1}-\mathbb{E}(V_{k-1}))^2 \\ n^{-3/2}M_k\mathbb{E}(X_{k-1}) \\ -n^{-1}M_k\,(V_{k-1}-\mathbb{E}(V_{k-1})) \\ n^{-1/2}M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \frac{1}{2}\sigma^2\int_0^1(\mathcal{B}_t-\widetilde{\mathcal{B}}_t)^2\,\mathrm{d}t \\ \frac{1}{2^{3/2}}\mu\sigma\left(\mathcal{B}_1+\widetilde{\mathcal{B}}_1-\int_0^1(\mathcal{B}_t+\widetilde{\mathcal{B}}_t)\,\mathrm{d}t\right) \\ \frac{1}{4}\sigma^2\left[(\mathcal{B}_1-\widetilde{\mathcal{B}}_1)^2-2\right] \\ \frac{1}{2^{1/2}}\sigma(\mathcal{B}_1+\widetilde{\mathcal{B}}_1) \end{bmatrix} \tag{8.2}$$

as $n \to \infty$, which yields the statement. Indeed, $\left(2^{-1/2}(\mathcal{B}_t+\widetilde{\mathcal{B}}_t)\right)_{t\in\mathbb{R}_+}$ and $\left(2^{-1/2}(\mathcal{B}_t-\widetilde{\mathcal{B}}_t)\right)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes, and by Itô's formula, $\int_0^1 t\,\mathrm{d}\widetilde{\mathcal{W}}_t = \widetilde{\mathcal{W}}_1 - \int_0^1 \widetilde{\mathcal{W}}_t\,\mathrm{d}t$ and $\int_0^1 \mathcal{W}_t\,\mathrm{d}\mathcal{W}_t = 2^{-1}(\mathcal{W}_1^2-1)$, which yield the statement with the choices $\widetilde{\mathcal{W}}_t := 2^{-1/2}(\mathcal{B}_t+\widetilde{\mathcal{B}}_t)$, $t \geq 0$, and $\mathcal{W}_t := 2^{-1/2}(\mathcal{B}_t-\widetilde{\mathcal{B}}_t)$, $t \geq 0$. Applying Lemmas B.2 and B.3 as in the proof of Proposition 3.1 and using Slutsky's lemma, (8.2) will follow from

$$\frac{1}{n^2}\sum_{k=1}^{n}(V_{k-1}-\mathbb{E}(V_{k-1}))^2 - \frac{1}{n}\sum_{k=1}^{\lfloor n/2\rfloor}\left(\mathcal{S}_{2k/n}^{(\lfloor n/2\rfloor)}-\mathcal{T}_{2k/n}^{(\lfloor n/2\rfloor)}\right)^2 \xrightarrow{\mathbb{P}} 0, \tag{8.3}$$

$$\frac{1}{n^{3/2}}\sum_{k=1}^{n}M_k\mathbb{E}(X_{k-1}) - \frac{\mu}{2^{3/2}}\left(\mathcal{S}_1^{(\lfloor n/2\rfloor)}+\mathcal{T}_1^{(\lfloor n/2\rfloor)}-\frac{2}{n}\sum_{k=1}^{\lfloor n/2\rfloor}\left(\mathcal{S}_{2k/n}^{(\lfloor n/2\rfloor)}+\mathcal{T}_{2k/n}^{(\lfloor n/2\rfloor)}\right)\right) \xrightarrow{\mathbb{P}} 0, \tag{8.4}$$

$$\frac{1}{n}\sum_{k=1}^{n}M_k\,(V_{k-1}-\mathbb{E}(V_{k-1})) + \frac{1}{4}\left[(\mathcal{S}_1^{(\lfloor n/2\rfloor)}-\mathcal{T}_1^{(\lfloor n/2\rfloor)})^2-2\sigma^2\right] \xrightarrow{\mathbb{P}} 0, \tag{8.5}$$

$$\frac{1}{n^{1/2}}\sum_{k=1}^{n}M_k - \frac{1}{2^{1/2}}\left(\mathcal{S}_1^{(\lfloor n/2\rfloor)}+\mathcal{T}_1^{(\lfloor n/2\rfloor)}\right) \xrightarrow{\mathbb{P}} 0. \tag{8.6}$$

We prove (8.3), (8.4), (8.5), and (8.6) only for the subsequence $(2n)_{n\in\mathbb{N}}$. Observe that

$$V_{2k}-\mathbb{E}(V_{2k}) = n^{1/2}(\mathcal{S}_{k/n}^{(n)}-\mathcal{T}_{k/n}^{(n)}), \qquad V_{2k-1}-\mathbb{E}(V_{2k-1}) = (\varepsilon_{2k}-\mu)-n^{1/2}(\mathcal{S}_{k/n}^{(n)}-\mathcal{T}_{k/n}^{(n)}),$$

for all $k \in \mathbb{N}$. Consequently,

$$\begin{aligned}\frac{1}{(2n)^2}\sum_{k=1}^{2n}(V_{k-1}-\mathbb{E}(V_{k-1}))^2 &= \frac{1}{2n}\sum_{k=1}^{n}(\mathcal{S}_{k/n}^{(n)}-\mathcal{T}_{k/n}^{(n)})^2 - \frac{1}{4n^2}(V_{2n}-\mathbb{E}(V_{2n}))^2 \\ &\quad - \frac{1}{2n^2}\sum_{k=1}^{n}(\varepsilon_{2k}-\mu)(V_{2k}-\mathbb{E}(V_{2k})) + \frac{1}{4n^2}\sum_{k=1}^{n}(\varepsilon_{2k}-\mu)^2.\end{aligned}$$

Thus, to prove (8.3) for the subsequence $(2n)_{n\in\mathbb{N}}$, it suffices to show that

$$\frac{1}{n^2}(V_{2n}-\mathbb{E}(V_{2n}))^2 \xrightarrow{\mathbb{P}} 0, \tag{8.7}$$

$$\frac{1}{n^2}\sum_{k=1}^{n}(\varepsilon_{2k}-\mu)(V_{2k}-\mathbb{E}(V_{2k})) \xrightarrow{\mathbb{P}} 0, \tag{8.8}$$

$$\frac{1}{n^2}\sum_{k=1}^{n}(\varepsilon_{2k}-\mu)^2 \xrightarrow{\mathbb{P}} 0, \tag{8.9}$$

as $n \to \infty$. We have $\mathbb{E}\left((V_{2n} - \mathbb{E}(V_{2n}))^2\right) = \mathrm{O}(n)$ and $\mathbb{E}\left((\varepsilon_{2k} - \mu)^2\right) = \sigma^2$; thus, we obtain (8.7) and (8.9). Further, $V_{2k} - \mathbb{E}(V_{2k}) = (\varepsilon_{2k} - \mu) - (V_{2k-1} - \mathbb{E}(V_{2k-1}))$; hence, (8.8) follows from (8.9) and from

$$\mathbb{E}\left(\left(\sum_{k=1}^{n} (\varepsilon_{2k} - \mu)(V_{2k-1} - \mathbb{E}(V_{2k-1}))\right)^2\right) = \sigma^2 \sum_{k=1}^{n} \mathbb{E}\left((V_{2k-1} - \mathbb{E}(V_{2k-1}))^2\right),$$

which is of order $\mathrm{O}(n^2)$; hence, we finished the proof of (8.3).

Now, we turn to prove (8.4). Observe that

$$\frac{1}{(2n)^{3/2}} \sum_{k=1}^{2n} M_k \mathbb{E}(X_{k-1}) = \frac{\mu}{2^{3/2}} \left(\mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)} - \frac{1}{n} \sum_{k=1}^{n} \left(\mathcal{S}_{k/n}^{(n)} + \mathcal{T}_{k/n}^{(n)}\right)\right) + \frac{\mu}{2^{3/2} n} \mathcal{S}_1^{(n)}.$$

Convergence (8.1) implies that $\mathcal{S}_1^{(n)} \xrightarrow{\mathcal{L}} \sigma \mathcal{B}_1$; thus, we obtain (8.4).

Now, we turn to prove (8.5). Observe that

$$\frac{1}{2n} \sum_{k=1}^{2n} M_k \left(V_{k-1} - \mathbb{E}(V_{k-1})\right) = -\frac{1}{4}\left[(\mathcal{S}_1^{(n)} - \mathcal{T}_1^{(n)})^2 - 2\sigma^2\right] + \frac{1}{4n} \sum_{k=1}^{2n} (\varepsilon_k - \mu)^2 - \frac{1}{2}\sigma^2.$$

By the strong law of large numbers, $(2n)^{-1} \sum_{k=1}^{2n} (\varepsilon_k - \mu)^2 \xrightarrow{\text{a.s.}} \sigma^2$ as $n \to \infty$; hence, we obtain (8.5).

Now, we turn to prove (8.6). First, observe that

$$\frac{1}{(2n)^{1/2}} \sum_{k=1}^{2n} M_k = \frac{1}{(2n)^{1/2}} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu) + \frac{1}{(2n)^{1/2}} \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu) = \frac{1}{2^{1/2}} \left(\mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)}\right);$$

thus, we obtain (8.6).

Finally, one can show (8.3), (8.4), (8.5), and (8.6) for the subsequence $(2n-1)_{n \in \mathbb{N}}$ in the same way. □

Appendix A: Estimations of moments

First, note that, for all $k \in \mathbb{N}$, $\mathbb{E}(M_k \mid \mathcal{F}_{k-1}) = 0$ and $\mathbb{E}(M_k) = 0$, because $M_k = X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1})$.

Lemma A.1. Let $(X_k)_{k \geq -1}$ be an INAR(2) process. Suppose that $X_0 = X_{-1} = 0$ and $\mathbb{E}(\varepsilon_1^2) < \infty$. Then, for all $k, \ell \in \mathbb{N}$ with $\ell \leq k$,

$$\begin{aligned}
\mathbb{E}(M_k M_\ell \mid \mathcal{F}_{k-1}) &= \begin{cases} \alpha(1-\alpha)X_{k-1} + \beta(1-\beta)X_{k-2} + \sigma^2 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell, \end{cases} \\
\mathbb{E}(M_k M_\ell) &= \begin{cases} \alpha(1-\alpha)\mathbb{E}(X_{k-1}) + \beta(1-\beta)\mathbb{E}(X_{k-2}) + \sigma^2 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell, \end{cases} \\
\mathbb{E}(M_k^3 \mid \mathcal{F}_{k-1}) &= \left[\mathbb{E}(\xi_{1,1} - \mathbb{E}(\xi_{1,1}))^3\right] X_{k-1} + \left[\mathbb{E}(\eta_{1,1} - \mathbb{E}(\eta_{1,1}))^3\right] X_{k-2} + \mathbb{E}(\varepsilon_1 - \mathbb{E}(\varepsilon_1))^3, \\
\mathbb{E}(M_k^3) &= \left[\mathbb{E}(\xi_{1,1} - \mathbb{E}(\xi_{1,1}))^3\right] \mathbb{E}(X_{k-1}) + \left[\mathbb{E}(\eta_{1,1} - \mathbb{E}(\eta_{1,1}))^3\right] \mathbb{E}(X_{k-2}) + \mathbb{E}(\varepsilon_1 - \mathbb{E}(\varepsilon_1))^3.
\end{aligned} \tag{A.1}$$

Proof. For all $k \in \mathbb{N}$, by (1.1) and (3.2),

$$M_k = \sum_{j=1}^{X_{k-1}} \left(\xi_{k,j} - \mathbb{E}(\xi_{k,j})\right) + \sum_{j=1}^{X_{k-2}} \left(\eta_{k,j} - \mathbb{E}(\eta_{k,j})\right) + (\varepsilon_k - \mathbb{E}(\varepsilon_k)) . \tag{A.2}$$

For each $k \in \mathbb{N}$, the random variables $\{\xi_{k,j} - \mathbb{E}(\xi_{k,j}),\ \eta_{k,j} - \mathbb{E}(\eta_{k,j}),\ \varepsilon_k - \mathbb{E}(\varepsilon_k) : j \in \mathbb{N}\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean; thus, we conclude the formulas for $\mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1})$ and $\mathbb{E}(M_k^2)$. If $\ell < k$, then $\mathbb{E}(M_k M_\ell \mid \mathcal{F}_{k-1}) = M_\ell \mathbb{E}(M_k \mid \mathcal{F}_{k-1}) = 0$. Thus, we obtain the formulas for $\mathbb{E}(M_k M_\ell \mid \mathcal{F}_{k-1})$ and $\mathbb{E}(M_k M_\ell)$ in case $k \neq \ell$. Multinomial theorem and (A.2) yield the formulas for $\mathbb{E}(M_k^3 \mid \mathcal{F}_{k-1})$ and $\mathbb{E}(M_k^3)$. □

The proof of the following Lemma is straightforward; see Lemma 9.2 in Barczy *et al.* (2012).

Lemma A.2. Let $(\zeta_k)_{k\in\mathbb{N}}$ be i.i.d. random variables such that $\mathbb{E}\left(|\zeta_1|^\ell\right) < \infty$ for some $\ell \in \mathbb{N}$.

((i)) If $\mathbb{E}(\zeta_1) \neq 0$, then there exists a polynomial Q_ℓ of degree ℓ such that its leading coefficient is $[\mathbb{E}(\zeta_1)]^\ell$ and

$$\mathbb{E}\left((\zeta_1 + \cdots + \zeta_N)^\ell\right) = Q_\ell(N), \qquad N \in \mathbb{N}.$$

((ii)) If $\mathbb{E}(\zeta_1) = 0$, then there exists a polynomial R_ℓ of degree at most $\ell/2$ such that

$$\mathbb{E}\left((\zeta_1 + \cdots + \zeta_N)^\ell\right) = R_\ell(N), \qquad N \in \mathbb{N}.$$

The coefficients of the polynomials Q_ℓ and R_ℓ depend on the moments $\mathbb{E}(\zeta_1^j)$, $j \in \{1, \ldots, \ell\}$.

Lemma A.3. Let $(X_k)_{k\geq -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in [0,1]^2$ such that $\alpha + \beta = 1$ (hence, it is unstable). Suppose $X_0 = X_{-1} = 0$ and $\mathbb{E}(\varepsilon_1^\ell) < \infty$ with some $\ell \in \mathbb{Z}_+$. Then there exists a constant c_ℓ such that $\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) \leq c_\ell n^\ell$, $n \in \mathbb{N}$, for all $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 \leq \ell$.

Proof. Observe that the statement is equivalent with the following: for each polynomial P of two variables having degree at most ℓ, there exists a constant c_P such that $\mathbb{E}\left(|P(X_n, X_{n-1})|\right) \leq c_P n^\ell$, $n \in \mathbb{N}$.

First, let us suppose that $(\alpha, \beta) \in (0,1)^2$. For $\ell = 0$, the statement is trivial. Let us suppose now that the statement holds for $0, 1, \ldots, \ell - 1$. Applying the multinomial theorem for $X_n^{\ell_1}$ and using that the random variables $\{\xi_{n,j}, \eta_{n,j}, \varepsilon_n : j \in \mathbb{N}\}$ are independent of each other and of the σ-algebra $\mathcal{F}_{n-1}$, we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2} \mid \mathcal{F}_{n-1})$$
$$= \sum_{\substack{k_1+k_2+k_3=\ell_1,\\ k_1,k_2,k_3\in\mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!k_3!} \mathbb{E}\left(\left(\sum_{j=1}^{M} \xi_{n,j}\right)^{k_1}\right)\Bigg|_{M=X_{n-1}} \mathbb{E}\left(\left(\sum_{j=1}^{N} \eta_{n,j}\right)^{k_2}\right)\Bigg|_{N=X_{n-2}} \mathbb{E}(\varepsilon_1^{k_3}) X_{n-1}^{\ell_2},$$

for all $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 = \ell$. Using part (i) of Lemma A.2 and separating the terms having degree ℓ and less than ℓ, we can write $\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2} \mid \mathcal{F}_{n-1})$ in the form

$$Q_{\ell_1,\ell_2}(X_{n-1}, X_{n-2}) + \sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} X_{n-1}^{k_1} \beta^{k_2} X_{n-2}^{k_2} X_{n-1}^{\ell_2},$$

where Q_{ℓ_1,ℓ_2} is a polynomial of two variables having degree at most $\ell-1$. Hence, $\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2})$ takes the form

$$\mathbb{E}\left(Q_{\ell_1,\ell_2}(X_{n-1},X_{n-2})\right)+\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\mathbb{E}\left(X_{n-1}^{\ell_2+k_1}X_{n-2}^{k_2}\right).$$

By the induction hypothesis (used for polynomials, see the beginning of the proof), there exists a constant $c_{Q_{\ell_1,\ell_2}}$ such that $\mathbb{E}\left(|Q_{\ell_1,\ell_2}(X_n,X_{n-1})|\right)\leq c_{Q_{\ell_1,\ell_2}}n^{\ell-1}$, $n\in\mathbb{N}$. In fact,

$$\mathbb{E}\left(|Q_{\ell_1,\ell_2}(X_n,X_{n-1})|\right)\leq\widetilde{c}_\ell n^{\ell-1} \tag{A.3}$$

for $n\in\mathbb{N}$ and $\ell_1,\ell_2\in\mathbb{Z}_+$ with $\ell_1+\ell_2=\ell$, where $\widetilde{c}_\ell:=\max_{0\leq i\leq\ell}c_{Q_{i,\ell-i}}$. Consequently,

$$\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2})\leq\widetilde{c}_\ell(n-1)^{\ell-1}+\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\mathbb{E}\left(X_{n-1}^{\ell_2+k_1}X_{n-2}^{k_2}\right).$$

Similarly, for all $k_1,k_2\in\mathbb{Z}_+$ with $k_1+k_2=\ell_1$, we have

$$\begin{aligned}\mathbb{E}\left(X_{n-1}^{\ell_2+k_1}X_{n-2}^{k_2}\right)&=\mathbb{E}\left(Q_{\ell_2+k_1,k_2}(X_{n-2},X_{n-3})\right)\\&\quad+\sum_{\substack{j_1+j_2=\ell_2+k_1,\\ j_1,j_2\in\mathbb{Z}_+}}\frac{(\ell_2+k_1)!}{j_1!j_2!}\alpha^{j_1}\beta^{j_2}\mathbb{E}\left(X_{n-2}^{k_2+j_1}X_{n-3}^{j_2}\right).\end{aligned}$$

Hence, we have

$$\begin{aligned}\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2})&=\mathbb{E}\left(Q_{\ell_1,\ell_2}(X_{n-1},X_{n-2})\right)\\&\quad+\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\sum_{\substack{j_1+j_2=\ell_2+k_1,\\ j_1,j_2\in\mathbb{Z}_+}}\frac{(\ell_2+k_1)!}{j_1!j_2!}\alpha^{j_1}\beta^{j_2}\mathbb{E}\left(X_{n-2}^{k_2+j_1}X_{n-3}^{j_2}\right)\\&\quad+\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\mathbb{E}\left(Q_{\ell_2+k_1,k_2}(X_{n-2},X_{n-3})\right).\end{aligned}$$

Applying (A.3) and

$$\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}=(\alpha+\beta)^{\ell_1}=1,$$

we conclude that

$$\begin{aligned}\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2})&\leq\widetilde{c}_\ell(n-1)^{\ell-1}+\widetilde{c}_\ell(n-2)^{\ell-1}\\&\quad+\sum_{\substack{k_1+k_2=\ell_1,\\ k_1,k_2\in\mathbb{Z}_+}}\frac{\ell_1!}{k_1!k_2!}\alpha^{k_1}\beta^{k_2}\sum_{\substack{j_1+j_2=\ell_2+k_1,\\ j_1,j_2\in\mathbb{Z}_+}}\frac{(\ell_2+k_1)!}{j_1!j_2!}\alpha^{j_1}\beta^{j_2}\mathbb{E}\left(X_{n-2}^{k_2+j_1}X_{n-3}^{j_2}\right).\end{aligned}$$

Using that $\mathbb{E}(X_1^rX_0^q)=0$, $r,q\in\mathbb{Z}_+$ (because $X_0=0$), after $n-1$ steps, we arrive at

$$\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2})\leq\widetilde{c}_\ell\sum_{i=1}^{n-1}i^{\ell-1}\leq\widetilde{c}_\ell n\cdot n^{\ell-1}=\widetilde{c}_\ell n^\ell,\qquad n\in\mathbb{N};$$

thus, the statement holds for all monomials of two variables having degree ℓ, and this implies the statement for all polynomials of two variables having degree ℓ.

Next, let us suppose that $(\alpha, \beta) = (1, 0)$. Then $X_n = X_{n-1} + \varepsilon_n$, $n \in \mathbb{N}$, which implies that $X_n = \sum_{i=1}^n \varepsilon_i$, $n \in \mathbb{N}$. By part (i) of Lemma A.2,

$$\mathbb{E}(X_n^\ell) = Q_\ell(n), \qquad n \in \mathbb{N}, \tag{A.4}$$

where Q_ℓ is a polynomial of degree ℓ. If $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 \leq \ell$, then using the independence of X_{n-1} and ε_n, we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \sum_{j=0}^{\ell_1} \binom{\ell_1}{j} \mathbb{E}(X_{n-1}^{j+\ell_2}) \mathbb{E}(\varepsilon_n^{\ell_1 - j}), \qquad n \in \mathbb{N}.$$

Using (A.4),

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \sum_{j=0}^{\ell_1} \binom{\ell_1}{j} Q_{j+\ell_2}(n-1) \mathbb{E}(\varepsilon_1^{\ell_1 - j}) = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N},$$

because for each $j \in \{0, \ldots, \ell_1\}$, the polynomial $Q_{j+\ell_2}$ is of degree $j + \ell_2 \leq \ell$, which yields the statement in case $(\alpha, \beta) = (1, 0)$.

Finally, let us suppose that $(\alpha, \beta) = (0, 1)$. Then $X_n = X_{n-2} + \varepsilon_n$, $n \in \mathbb{N}$, which implies that $X_{2n} = \sum_{i=1}^n \varepsilon_{2i}$, $X_{2n-1} = \sum_{i=1}^n \varepsilon_{2i-1}$, $n \in \mathbb{N}$. By part (i) of Lemma A.2, we have $\mathbb{E}(X_{2n}^\ell) = \mathbb{E}(X_{2n-1}^\ell) = Q_\ell(n)$, $n \in \mathbb{N}$, where Q_ℓ is a polynomial of degree ℓ. Using the independence of X_{2n} and X_{2n-1}, for $\ell_1 + \ell_2 \leq \ell$, $\ell_1, \ell_2 \in \mathbb{Z}_+$, we have

$$\mathbb{E}(X_{2n}^{\ell_1} X_{2n-1}^{\ell_2}) = \mathbb{E}(X_{2n}^{\ell_1}) \mathbb{E}(X_{2n-1}^{\ell_2}) = Q_{\ell_1}(n) Q_{\ell_2}(n) = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N},$$

as desired. The expectation $\mathbb{E}(X_{2n-1}^{\ell_1} X_{2n-2}^{\ell_2})$ can be handled in a similar way.

On page 46 in Barczy *et al.* (2012), one can find another proof of this lemma. □

Corollary A.4. Let $(X_k)_{k \geq -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in [0, 1]^2$ such that $\alpha + \beta = 1$ (hence, it is unstable). Suppose that $X_0 = X_{-1} = 0$ and $\mathbb{E}(\varepsilon_1^\ell) < \infty$ with some $\ell \in \mathbb{N}$. Then

$$\mathbb{E}(X_k^i) = \mathrm{O}(k^i), \quad \mathbb{E}(M_k^i) = \mathrm{O}(k^{\lfloor i/2 \rfloor}), \quad \mathbb{E}(U_k^i) = \mathrm{O}(k^i), \quad \mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$$

for $k \in \mathbb{N}$ and $i, j \in \mathbb{Z}_+$ with $i \leq \ell$ and $2j \leq \ell$.

Proof. The estimate $\mathbb{E}(X_k^i) = \mathrm{O}(k^i)$ readily follows by Lemma A.3. Next, we turn to prove $\mathbb{E}(M_k^i) = \mathrm{O}(k^{\lfloor i/2 \rfloor})$. Using (A.2) and that the random variables $\{\xi_{k,j}, \eta_{k,j}, \varepsilon_k : j \in \mathbb{N}\}$ are independent of each other and of the σ-algebra $\mathcal{F}_{k-1}$, we have for all $k \in \mathbb{N}$,

$$\mathbb{E}(M_k^i \mid \mathcal{F}_{k-1}) = \sum_{\substack{i_1+i_2+i_3=i, \\ i_1, i_2, i_3 \in \mathbb{Z}_+}} \frac{i!}{i_1! i_2! i_3!} \mathbb{E}\left(\left(\sum_{j=1}^M (\xi_{k,j} - \mathbb{E}(\xi_{k,j})) \right)^{i_1} \right) \Bigg|_{M = X_{k-1}}$$

$$\times \mathbb{E}\left(\left(\sum_{j=1}^N (\eta_{k,j} - \mathbb{E}(\eta_{k,j})) \right)^{i_2} \right) \Bigg|_{N = X_{k-2}} \mathbb{E}\left((\varepsilon_k - \mathbb{E}(\varepsilon_k))^{i_3} \right).$$

By part (ii) of Lemma A.2, there exist polynomials Q_{i_1}, $i_1 \in \mathbb{N}$, of degree at most $i_1/2$, and $\widetilde{Q}_{i_2}$, $i_2 \in \mathbb{N}$, of degree at most $i_2/2$ such that

$$\mathbb{E}(M_k^i \mid \mathcal{F}_{k-1}) = \sum_{\substack{i_1+i_2+i_3=i,\\ i_1,i_2,i_3\in\mathbb{Z}_+}} \frac{i!}{i_1!i_2!i_3!} Q_{i_1}(X_{k-1})\widetilde{Q}_{i_2}(X_{k-2})\mathbb{E}\left((\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3}\right).$$

Hence,

$$\mathbb{E}(M_k^i) = \sum_{\substack{i_1+i_2+i_3=i,\\ i_1,i_2,i_3\in\mathbb{Z}_+}} \frac{i!}{i_1!i_2!i_3!} \mathbb{E}\left(Q_{i_1}(X_{k-1})\widetilde{Q}_{i_2}(X_{k-2})\right)\mathbb{E}\left((\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3}\right)$$

for $k \in \mathbb{N}$. Clearly, $Q_{i_1}(X_{k-1})\widetilde{Q}_{i_2}(X_{k-2}) = Q^*_{i_1+i_2}(X_{k-1}, X_{k-2})$, where $Q^*_{i_1+i_2}$ is a polynomial of two variables having degree at most $(i_1 + i_2)/2 \le i/2$, and hence at most $\lfloor i/2 \rfloor$. By Lemma A.3, there exists a constant $c_{Q^*_{i_1+i_2}}$ such that $\mathbb{E}(|Q^*_{i_1+i_2}(X_{k-1}, X_{k-2})|) \le c_{Q^*_{i_1+i_2}}(k-1)^{\lfloor i/2 \rfloor}$. Hence

$$|\mathbb{E}(M_k^i)| \le (k-1)^{\lfloor i/2 \rfloor} \sum_{\substack{i_1+i_2+i_3=i,\\ i_1,i_2,i_3\in\mathbb{Z}_+}} \frac{i!}{i_1!i_2!i_3!} c_{Q^*_{i_1+i_2}} \left|\mathbb{E}\left((\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3}\right)\right|,$$

for all $k \in \mathbb{N}$, as desired.

Next, we turn to prove $\mathbb{E}(U_k^i) = \mathrm{O}(k^i)$, $i, k \in \mathbb{N}$ with $i \le \ell$. First, note that $(a+b)^i \le 2^{i-1}(a^i + b^i)$, $a, b \ge 0$. Hence, by Lemma A.3,

$$\mathbb{E}(U_k^i) = \mathbb{E}((X_k + \beta X_{k-1})^i) \le 2^{i-1}(\mathbb{E}(X_k^i) + \beta^i \mathbb{E}(X_{k-1}^i)) \le c_i 2^i k^i.$$

Finally, for $2j \le \ell$, $j \in \mathbb{Z}_+$, we prove $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, using induction in k. By the recursion (3.9), we have $\mathbb{E}(V_k) = -\beta\mathbb{E}(V_{k-1}) + \mu$, $k \in \mathbb{N}$, with initial value $\mathbb{E}(V_0) = 0$; hence, $\mathbb{E}(V_k) = \mu\sum_{i=0}^{k-1}(-\beta)^i$, $k \in \mathbb{N}$, which yields that $|\mathbb{E}(V_k)| = \mathrm{O}(1)$. Let us introduce the notation $\widetilde{V}_k := V_k - \mathbb{E}(V_k)$, $k \in \mathbb{N}$. Because, by the triangular inequality for the L_{2j}-norm,

$$\left(\mathbb{E}(V_k^{2j})\right)^{\frac{1}{2j}} \le \left(\mathbb{E}(\widetilde{V}_k^{2j})\right)^{\frac{1}{2j}} + |\mathbb{E}(V_k)|,$$

and $|\mathbb{E}(V_k)| = \mathrm{O}(1)$; for proving $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, it is enough to show that $\mathbb{E}(\widetilde{V}_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$. Using again the recursion (3.9), we obtain $\widetilde{V}_k = -\beta\widetilde{V}_{k-1} + M_k$, $k \in \mathbb{N}$. Hence, by the induction hypothesis,

$$\left(\mathbb{E}(\widetilde{V}_k^{2j})\right)^{\frac{1}{2j}} \le \beta\left(\mathbb{E}(\widetilde{V}_{k-1}^{2j})\right)^{\frac{1}{2j}} + \left(\mathbb{E}(M_k^{2j})\right)^{\frac{1}{2j}} = \left(\mathrm{O}(k^j)\right)^{\frac{1}{2j}},$$

thus, $\mathbb{E}(\widetilde{V}_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, as desired. □

Corollary A.5. Let $(X_k)_{k\ge -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in [0,1]^2$ such that $\alpha + \beta = 1$ (hence, it is unstable). Suppose that $X_0 = X_{-1} = 0$ and $\mathbb{E}(\varepsilon_1^\ell) < \infty$ with some $\ell \in \mathbb{N}$. Then

(i) for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \le \ell/2$, and for all $\kappa > i + \frac{j}{2} + 1$, we have

$$n^{-\kappa}\sum_{k=1}^n |U_k^i V_k^j| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty, \tag{A.5}$$

(ii) for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \le \ell$, for all $T > 0$, and for all $\kappa > i + \frac{j}{2} + \frac{i+j}{\ell}$, we have

$$n^{-\kappa} \sup_{t\in[0,T]} |U^i_{\lfloor nt\rfloor} V^j_{\lfloor nt\rfloor}| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty, \tag{A.6}$$

(iii) for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \le \ell/4$, for all $T > 0$, and for all $\kappa > i + \frac{j}{2} + \frac{1}{2}$, we have

$$n^{-\kappa} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt\rfloor} [U^i_k V^j_k - \mathbb{E}(U^i_k V^j_k \mid \mathcal{F}_{k-1})] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty. \tag{A.7}$$

Proof. The statements are easy consequences of Cauchy–Schwarz's inequality, Markov's inequality, Doob's maximal inequality, and Corollary A.4; see Corollary 9.2 in Barczy *et al.* (2012). □

Appendix B: A version of the continuous mapping theorem

A function $f : \mathbb{R}_+ \to \mathbb{R}^d$ is called *càdlàg* if it is right continuous with left limits. Let $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg and continuous functions on $\mathbb{R}_+$, respectively. Let $\mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ denote the Borel σ-algebra on $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ for the metric defined in Chapter VI, (1.26) of Jacod & Shiryaev (2003). With this metric, $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ is a complete and separable metric space, and the topology induced by this metric is the so-called Skorokhod topology. For $\mathbb{R}^d$-valued stochastic processes $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ and $(\mathcal{Y}^{(n)}_t)_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, with càdlàg paths, we write $\mathcal{Y}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{Y}$ if the distribution of $\mathcal{Y}^{(n)}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ converges weakly to the distribution of $\mathcal{Y}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ as $n \to \infty$. Concerning the notation $\xrightarrow{\mathcal{L}}$, we note that if ξ and ξ_n, $n \in \mathbb{N}$, are random elements with values in a metric space (E, d), then we also denote by $\xi_n \xrightarrow{\mathcal{L}} \xi$ the weak convergence of the distributions of ξ_n on the space $(E, \mathcal{B}(E))$ toward the distribution of ξ on the space $(E, \mathcal{B}(E))$ as $n \to \infty$, where $\mathcal{B}(E)$ denotes the Borel σ-algebra on E induced by the given metric d.

The following version of the continuous mapping theorem can be found for example in Theorem 3.27 of Kallenberg (1997).

Lemma B.1. Let (S, d_S) and (T, d_T) be metric spaces and $(\xi_n)_{n\in\mathbb{N}}$, ξ be random elements with values in S such that $\xi_n \xrightarrow{\mathcal{L}} \xi$ as $n \to \infty$. Let $f : S \to T$ and $f_n : S \to T$, $n \in \mathbb{N}$, be measurable mappings and $C \in \mathcal{B}(S)$ such that $\mathbb{P}(\xi \in C) = 1$ and $\lim_{n\to\infty} d_T(f_n(s_n), f(s)) = 0$ if $\lim_{n\to\infty} d_S(s_n, s) = 0$ and $s \in C$. Then $f_n(\xi_n) \xrightarrow{\mathcal{L}} f(\xi)$ as $n \to \infty$.

For the case $S := \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $T := \mathbb{R}^q$, where $d, q \in \mathbb{N}$, we formulate a consequence of Lemma B.1. For functions f and f_n, $n \in \mathbb{N}$, in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, we write $f_n \xrightarrow{\text{lu}} f$ if $(f_n)_{n\in\mathbb{N}}$ converges to f locally uniformly, that is, if $\sup_{t\in[0,T]} \|f_n(t) - f(t)\| \to 0$ as $n \to \infty$ for all $T > 0$. For measurable mappings $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$, $n \in \mathbb{N}$, we will denote by $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ the set of all functions $f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that $\Phi_n(f_n) \to \Phi(f)$ whenever $f_n \xrightarrow{\text{lu}} f$ with $f_n \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$. We will use the following version of the continuous mapping theorem several times; see, for example, Lemma 3.1 of Ispány & Pap (2010).

Lemma B.2. Let $d, q \in \mathbb{N}$, and let $(\mathcal{U}_t)_{t\in\mathbb{R}_+}$, $(\mathcal{U}^{(n)}_t)_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, be $\mathbb{R}^d$-valued stochastic processes with càdlàg paths such that $\mathcal{U}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{U}$. Let $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ and Φ_n :

$\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$, $n \in \mathbb{N}$, be measurable mappings such that there exists $C \subset C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ with $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ and $\mathbb{P}(\boldsymbol{\mathcal{U}} \in C) = 1$. Then $\Phi_n(\boldsymbol{\mathcal{U}}^{(n)}) \xrightarrow{\mathcal{L}} \Phi(\boldsymbol{\mathcal{U}})$ as $n \to \infty$.

In order to apply Lemma B.2, we will use the following statement several times; see Lemma B.3 in Barczy *et al.* (2012).

Lemma B.3. Let $d, p, q \in \mathbb{N}$, $h : \mathbb{R}^d \to \mathbb{R}^q$ be a continuous function and $K : [0, 1] \times \mathbb{R}^{2d} \to \mathbb{R}^p$ be a function such that for all $R > 0$, there exists a constant $C_R > 0$ such that

$$\|K(s, x) - K(t, y)\| \le C_R(|t - s| + \|x - y\|),$$

for all $s, t \in [0, 1]$ and $x, y \in \mathbb{R}^{2d}$ with $\|x\| \le R$ and $\|y\| \le R$. Moreover, let us define the mappings $\Phi, \Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^{q+p}$, $n \in \mathbb{N}$, by

$$\Phi_n(f) := \left(h(f(1)), \frac{1}{n}\sum_{k=1}^{n} K\left(\frac{k}{n}, f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right)\right)\right),$$

$$\Phi(f) := \left(h(f(1)), \int_0^1 K(u, f(u), f(u))\,\mathrm{d}u\right),$$

for all $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$. Then the mappings Φ and Φ_n, $n \in \mathbb{N}$, are measurable, and $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$.

Appendix C: Convergence of random step processes

We recall a result about convergence of random step processes toward a diffusion process; see Ispány & Pap (2010). This result is used for the proof of convergence (5.4).

Theorem C.1. Let $\boldsymbol{\gamma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ be a continuous function. Assume that uniqueness in the sense of probability law holds for the SDE

$$\mathrm{d}\boldsymbol{\mathcal{U}}_t = \gamma(t, \boldsymbol{\mathcal{U}}_t)\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \tag{C.1}$$

with initial value $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{u}_0$ for all $\boldsymbol{u}_0 \in \mathbb{R}^d$, where $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ is an r-dimensional standard Wiener process. Let $(\boldsymbol{\mathcal{U}}_t)_{t\in\mathbb{R}_+}$ be a solution of (C.1) with initial value $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{0} \in \mathbb{R}^d$.

For each $n \in \mathbb{N}$, let $(\boldsymbol{U}_k^{(n)})_{k\in\mathbb{N}}$ be a sequence of d-dimensional martingale differences with respect to a filtration $(\mathcal{F}_k^{(n)})_{k\in\mathbb{Z}_+}$ with $\mathbb{E}\left(\|\boldsymbol{U}_k^{(n)}\|^2\right) < \infty$ for all $k \in \mathbb{N}$. Let

$$\boldsymbol{\mathcal{U}}_t^{(n)} := \sum_{k=1}^{\lfloor nt\rfloor} \boldsymbol{U}_k^{(n)}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Suppose that, for each $T > 0$,

(i) $\sup\limits_{t\in[0,T]} \left\| \sum\limits_{k=1}^{\lfloor nt\rfloor} \mathbb{E}\left(\boldsymbol{U}_k^{(n)}(\boldsymbol{U}_k^{(n)})^\top \,\middle|\, \mathcal{F}_{k-1}^{(n)}\right) - \int_0^t \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})\boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})^\top \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0,$

(ii) $\sum\limits_{k=1}^{\lfloor nT\rfloor} \mathbb{E}\left(\|\boldsymbol{U}_k^{(n)}\|^2 \;{}_{\{\|\boldsymbol{U}_k^{(n)}\|>\theta\}} \,\middle|\, \mathcal{F}_{k-1}^{(n)}\right) \xrightarrow{\mathbb{P}} 0$ for all $\theta > 0$.

Then $\boldsymbol{\mathcal{U}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{U}}$ as $n \to \infty$.

Acknowledgements

The research of M. Barczy and G. Pap was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program – Elaborating and operating an inland student and researcher personal support system'. The project was subsidized by the European Union and co-financed by the European Social Fund. M. Ispány has been supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

References

Alzaid, A. A. & Al-Osh, M. A. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Probab.* **27**, 314–324.

Barczy, M., Ispány, M & Pap, G. (2010). *Asymptotic behavior of CLS estimators of autoregressive parameter for nonprimitive unstable INAR(2) models*. Available on the ArXiv: http://arxiv.org/abs/1006.4641.

Barczy, M., Ispány, M. & Pap, G. (2011). Asymptotic behavior of unstable INAR(p) processes. *Stochastic Process. Appl.* **121**, 583–608.

Barczy, M., Ispány, M. & Pap, G. (2012). *Asymptotic behavior of CLS estimators for unstable INAR(2) models*. Available on the ArXiv: http://arxiv.org/abs/1202.1617.

Brännäs, K. & Hellström, J. (2001). Generalized integer-valued autoregression. *Econometric Rev.* **20**, 425–443.

Drost, F. C., van den Akker, R. & Werker, B. J. M. (2008). Local asymptotic normality and efficient estimation for INAR(p) models. *J. Time Series Anal.* **29**, 783–801.

Drost, F. C., van den Akker, R. & Werker, B. J. M. (2009). The asymptotic structure of nearly unstable non-negative integer-valued AR(1) models. *Bernoulli* **15**, 297–324.

Du, J. G. & Li, Y. (1991). The integer valued autoregressive (INAR(p)) model. *J. Time Series Anal.* **12**, 129–142.

Franke, J. & Seligmann, T. (1993). Conditional maximum-likelihood estimates for INAR(1) processes and their applications to modelling epileptic seizure counts. In *Developments in time series analysis* (ed Subba Rao, T.), Chapman & Hall, London; 310–330.

Gauthier, G. & Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR(p). *Ann. Sci. Math. Québec* **18**, 49–71.

Hall, P. & Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71**, 285–317.

Hamilton, J. D. (1994). *Time series analysis*, Princeton University Press, Princeton.

Ispány, M., Pap, G. & van Zuijlen, M. C. A. (2003a). Asymptotic behaviour of estimators of the parameters of nearly unstable INAR(1) models. In *Foundations of Statistical Inference (Shoresh, 2000)* Physica, Heidelberg; 193–204.

Ispány, M., Pap, G. & van Zuijlen, M. C. A. (2003b). Asymptotic inference for nearly unstable INAR(1) models. *J. Appl. Probab.* **40**, 750–765.

Ispány, M. & Pap, G. (2010). A note on weak convergence of random step processes. *Acta Math. Hungar.* **126**, 381–395.

Jacod, J. & Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*, (2nd ed.)., Springer-Verlag, Berlin.

Kallenberg, O. (1997). *Foundations of modern probability*, Springer, New York, Berlin, Heidelberg.

Kesten, H. & Stigum, B. P. (1966a). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.

Kesten, H. & Stigum, B. P. (1966b). Additional limit theorems for indecomposable multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1463–1481.

Kesten, H. & Stigum, B. P. (1967). Limit theorems for decomposable multi-dimensional Galton–Watson processes. *J. Math. Anal. Appl.* **17**, 309–338.

Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive processes. *J. Time Series Anal.* **19**, 439–455.

Mikosch, T. & Straumann, D. (2002). Whittle estimation in a heavy-tailed GARCH(1,1) model. *Stochastic Process. Appl.* **100**, 187–222.

Revuz, D. & Yor, M. (2001). *Continuous martingales and Brownian motion*, (3rd ed.)., Corrected 2nd Printing, Springer-Verlag, Berlin.
Sims, C. A., Stock, J. H. & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica* **58**, 113–144.
Steutel, F. & van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.* **7**, 893–899.
Tanaka, K. (1996). *Time series analysis, nonstationary and noninvertible distribution theory*, John Wiley & Sons, Inc., New York.
Weiß, C. H. (2008). Thinning operations for modelling time series of counts—a survey. *AStA Adv. Stat. Anal.* **92**, 319–341.

Received 16 December 2012, in final form 29 October 2013

Mátyás Barczy, Applied Mathematics and Probability, Faculty of Informatics. AQ3
E-mail: barczy.matyas@inf.unideb.hu